GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:32:42 ; Search time 78.5 Seconds
(without alignments)
27.986 Million cell updates/sec

Title: US-10-771-242-293
Perfect score: 9
Sequence: 1 RXXXX 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5	55.6	2	2	ADH29628	Adh29628 Swinepox
2	5	55.6	2	4	AAM98426	Aam98426 Human pep
3	5	55.6	2	4	AAB91738	Aab91738 Opioid pe
4	5	55.6	2	5	ABG93548	Abg93548 Human P-g
5	5	55.6	2	5	ABG93543	Abg93543 Human P-g
6	5	55.6	2	5	ABG93498	Abg93498 Human P-g
7	5	55.6	2	5	ABG93615	Abg93615 Human P-g
8	5	55.6	2	5	ABG93621	Abg93621 Human P-g
9	5	55.6	2	5	ABG93576	Abg93576 Human P-g
10	5	55.6	2	5	ABG93550	Abg93550 Human P-g
11	5	55.6	2	5	ABP53564	Abp53564 Plasmid p
12	5	55.6	2	6	ABR39485	Abr39485 AlphaS1 c
13	5	55.6	2	7	AAE39209	Aae39209 Angiogene
14	5	55.6	2	7	ADL98393	Adl98393 Human leu
15	5	55.6	2	7	ADW36641	Adw36641 HLA bindi
16	5	55.6	2	7	ADW36964	Adw36964 HLA bindi
17	5	55.6	2	7	ADW37285	Adw37285 HLA bindi
18	5	55.6	3	1	AAP30601	Aap30601 Sequence
19	5	55.6	3	1	AAP90668	Aap90668 New antih
20	5	55.6	3	1	AAP90665	Aap90665 New antih
21	5	55.6	3	1	AAP97811	Aap97811 Sequence
22	5	55.6	3	2	AAR00718	Aar00718 Core repe
23	5	55.6	3	2	AAR04607	Aar04607 Antiviral
24	5	55.6	3	2	AAR10543	Aar10543 Hypotensi
25	5	55.6	3	2	AAR32271	Aar32271 Soybean g
26	5	55.6	3	2	AAR36707	Aar36707 Adhesion
27	5	55.6	3	2	AAR53144	Aar53144 RGD pepti
28	5	55.6	3	2	AAR38487	Aar38487 Human RDS
29	5	55.6	3	2	AAR42569	Aar42569 Peptide c
30	5	55.6	3	2	AAR30753	Aar30753 IgE-mast
31	5	55.6	3	2	AAR48960	Aar48960 NL4-3 tru
32	5	55.6	3	2	AAR46824	Aar46824 Phytase d
33	5	55.6	3	2	AAR44666	Aar44666 Platelet
34	5	55.6	3	2	AAR63264	Aar63264 Thrombin
35	5	55.6	3	2	AAR63263	Aar63263 Thrombin
36	5	55.6	3	2	AAR63261	Aar63261 Thrombin
37	5	55.6	3	2	AAR63260	Aar63260 Thrombin
38	5	55.6	3	2	AAR63262	Aar63262 Thrombin
39	5	55.6	3	2	AAR51440	Aar51440 IGF-1 ana
40	5	55.6	3	2	AAR61094	Aar61094 ACE-inhib
41	5	55.6	3	2	AAR61092	Aar61092 ACE-inhib
42	5	55.6	3	2	AAR61093	Aar61093 ACE-inhib
43	5	55.6	3	2	AAR73934	Aar73934 Novel tri
44	5	55.6	3	2	AAR57458	Aar57458 Lactoferr
45	5	55.6	3	2	AAR57450	Aar57450 Lactoferr
46	5	55.6	3	2	AAR48523	Aar48523 Lactoferr
47	5	55.6	3	2	AAR48527	Aar48527 Lactoferr
48	5	55.6	3	2	AAR48519	Aar48519 Lactoferr
49	5	55.6	3	2	AAR58575	Aar58575 Angiotens
50	5	55.6	3	2	AAR58571	Aar58571 Angiotens
51	5	55.6	3	2	AAR58576	Aar58576 Angiotens
52	5	55.6	3	2	AAR58578	Aar58578 Angiotens
53	5	55.6	3	2	AAR58573	Aar58573 Angiotens
54	5	55.6	3	2	AAR58569	Aar58569 Angiotens
55	5	55.6	3	2	AAR58570	Aar58570 Angiotens
56	5	55.6	3	2	AAR58572	Aar58572 Angiotens
57	5	55.6	3	2	AAR58577	Aar58577 Angiotens
58	5	55.6	3	2	AAR84695	Aar84695 Bovine la
59	5	55.6	3	2	AAR84687	Aar84687 Bovine la
60	5	55.6	3	2	AAR84691	Aar84691 Bovine la
61	5	55.6	3	2	AAR69779	Aar69779 Thrombosp
62	5	55.6	3	2	AAR70472	Aar70472 Cancer me
63	5	55.6	3	2	AAR82919	Aar82919 Non-RGD,
64	5	55.6	3	2	AAR85069	Aar85069 Calcium i
65	5	55.6	3	2	AAR64657	Aar64657 HPP3 pept
66	5	55.6	3	2	AAR64722	Aar64722 HPP3 pept
67	5	55.6	3	2	AAR90597	Aar90597 Lactoferr
68	5	55.6	3	2	AAR90605	Aar90605 Lactoferr
69	5	55.6	3	2	AAR90601	Aar90601 Lactoferr
70	5	55.6	3	2	AAR99827	Aar99827 Mutant ch
71	5	55.6	3	2	AAR99828	Aar99828 Active do
72	5	55.6	3	2	AAR98543	Aar98543 Peptide f
73	5	55.6	3	2	AAR98547	Aar98547 Peptide f
74	5	55.6	3	2	AAR98551	Aar98551 Peptide f
75	5	55.6	3	2	AAR88897	Aar88897 Small syn
76	5	55.6	3	2	AAR88899	Aar88899 Small syn
77	5	55.6	3	2	AAR88900	Aar88900 Small syn
78	5	55.6	3	2	AAR88898	Aar88898 Small syn
79	5	55.6	3	2	AAW11094	Aaw11094 Platelet-
80	5	55.6	3	2	AAR89785	Aar89785 Cell adhe
81	5	55.6	3	2	AAW00269	Aaw00269 Cytokine
82	5	55.6	3	2	AAW25173	Aaw25173 RGD-pepti
83	5	55.6	3	2	AAW18871	Aaw18871 Peptide f
84	5	55.6	3	2	AAW33241	Aaw33241 Analgesic
85	5	55.6	3	2	AAW31314	Aaw31314 Human Int
86	5	55.6	3	2	AAW31143	Aaw31143 Platelet-
87	5	55.6	3	2	AAW56177	Aaw56177 Anti-infl
88	5	55.6	3	2	AAW56181	Aaw56181 Anti-infl
89	5	55.6	3	2	AAW56183	Aaw56183 Anti-infl
90	5	55.6	3	2	AAW56186	Aaw56186 Anti-infl
91	5	55.6	3	2	AAW56196	Aaw56196 Anti-infl
92	5	55.6	3	2	AAW56201	Aaw56201 Anti-infl
93	5	55.6	3	2	AAW56234	Aaw56234 Anti-infl
94	5	55.6	3	2	AAW56200	Aaw56200 Anti-infl
95	5	55.6	3	2	AAW56209	Aaw56209 Anti-infl
96	5	55.6	3	2	AAW56233	Aaw56233 Anti-infl
97	5	55.6	3	2	AAW56188	Aaw56188 Anti-infl

98	5	55.6	3	2	AAW56197	Aaw56197	Anti-infl
99	5	55.6	3	2	AAW56207	Aaw56207	Anti-infl
100	5	55.6	3	2	AAW56220	Aaw56220	Anti-infl
101	5	55.6	3	2	AAW56239	Aaw56239	Anti-infl
102	5	55.6	3	2	AAW56175	Aaw56175	Anti-infl
103	5	55.6	3	2	AAW56190	Aaw56190	Anti-infl
104	5	55.6	3	2	AAW56195	Aaw56195	Anti-infl
105	5	55.6	3	2	AAW56230	Aaw56230	Anti-infl
106	5	55.6	3	2	AAW56246	Aaw56246	Anti-infl
107	5	55.6	3	2	AAW56185	Aaw56185	Anti-infl
108	5	55.6	3	2	AAW56191	Aaw56191	Anti-infl
109	5	55.6	3	2	AAW56193	Aaw56193	Anti-infl
110	5	55.6	3	2	AAW56198	Aaw56198	Anti-infl
111	5	55.6	3	2	AAW56213	Aaw56213	Anti-infl
112	5	55.6	3	2	AAW56229	Aaw56229	Anti-infl
113	5	55.6	3	2	AAW56240	Aaw56240	Anti-infl
114	5	55.6	3	2	AAW56178	Aaw56178	Anti-infl
115	5	55.6	3	2	AAW56202	Aaw56202	Anti-infl
116	5	55.6	3	2	AAW56212	Aaw56212	Anti-infl
117	5	55.6	3	2	AAW56216	Aaw56216	Anti-infl
118	5	55.6	3	2	AAW56232	Aaw56232	Anti-infl
119	5	55.6	3	2	AAW56241	Aaw56241	Anti-infl
120	5	55.6	3	2	AAW56172	Aaw56172	Anti-infl
121	5	55.6	3	2	AAW56173	Aaw56173	Anti-infl
122	5	55.6	3	2	AAW56180	Aaw56180	Anti-infl
123	5	55.6	3	2	AAW56182	Aaw56182	Anti-infl
124	5	55.6	3	2	AAW56203	Aaw56203	Anti-infl
125	5	55.6	3	2	AAW56204	Aaw56204	Anti-infl
126	5	55.6	3	2	AAW56208	Aaw56208	Anti-infl
127	5	55.6	3	2	AAW56210	Aaw56210	Anti-infl
128	5	55.6	3	2	AAW56221	Aaw56221	Anti-infl
129	5	55.6	3	2	AAW56236	Aaw56236	Anti-infl
130	5	55.6	3	2	AAW56206	Aaw56206	Anti-infl
131	5	55.6	3	2	AAW56214	Aaw56214	Anti-infl
132	5	55.6	3	2	AAW56218	Aaw56218	Anti-infl
133	5	55.6	3	2	AAW56222	Aaw56222	Anti-infl
134	5	55.6	3	2	AAW56235	Aaw56235	Anti-infl
135	5	55.6	3	2	AAW56238	Aaw56238	Anti-infl
136	5	55.6	3	2	AAW56248	Aaw56248	Anti-infl
137	5	55.6	3	2	AAW56187	Aaw56187	Anti-infl
138	5	55.6	3	2	AAW56227	Aaw56227	Anti-infl
139	5	55.6	3	2	AAW56194	Aaw56194	Anti-infl
140	5	55.6	3	2	AAW56199	Aaw56199	Anti-infl
141	5	55.6	3	2	AAW56211	Aaw56211	Anti-infl
142	5	55.6	3	2	AAW56245	Aaw56245	Anti-infl
143	5	55.6	3	2	AAW56189	Aaw56189	Anti-infl
144	5	55.6	3	2	AAW56219	Aaw56219	Anti-infl
145	5	55.6	3	2	AAW56174	Aaw56174	Anti-infl
146	5	55.6	3	2	AAW56205	Aaw56205	Anti-infl
147	5	55.6	3	2	AAW56217	Aaw56217	Anti-infl
148	5	55.6	3	2	AAW56223	Aaw56223	Anti-infl
149	5	55.6	3	2	AAW56224	Aaw56224	Anti-infl
150	5	55.6	3	2	AAW56231	Aaw56231	Anti-infl
151	5	55.6	3	2	AAW56244	Aaw56244	Anti-infl
152	5	55.6	3	2	AAW56179	Aaw56179	Anti-infl
153	5	55.6	3	2	AAW56237	Aaw56237	Anti-infl
154	5	55.6	3	2	AAW56242	Aaw56242	Anti-infl
155	5	55.6	3	2	AAW56243	Aaw56243	Anti-infl
156	5	55.6	3	2	AAW56176	Aaw56176	Anti-infl
157	5	55.6	3	2	AAW56247	Aaw56247	Anti-infl
158	5	55.6	3	2	AAW56171	Aaw56171	Anti-infl
159	5	55.6	3	2	AAW56184	Aaw56184	Anti-infl
160	5	55.6	3	2	AAW56192	Aaw56192	Anti-infl
161	5	55.6	3	2	AAW56215	Aaw56215	Anti-infl
162	5	55.6	3	2	AAW56225	Aaw56225	Anti-infl
163	5	55.6	3	2	AAW56226	Aaw56226	Anti-infl
164	5	55.6	3	2	AAW56228	Aaw56228	Anti-infl
165	5	55.6	3	2	AAW64738	Aaw64738	Angiotens
166	5	55.6	3	2	AAW48589	Aaw48589	Integrin
167	5	55.6	3	2	AAW52450	Aaw52450	Loop regi
168	5	55.6	3	2	AAW52449	Aaw52449	Loop regi
169	5	55.6	3	2	AAW52447	Aaw52447	Loop regi
170	5	55.6	3	2	AAW61942	Aaw61942	PPI bindi
171	5	55.6	3	2	AAW63167	Aaw63167	Mouse mas
172	5	55.6	3	2	AAW63169	Aaw63169	Mouse mas
173	5	55.6	3	2	AAW63161	Aaw63161	Mouse mas
174	5	55.6	3	2	AAW63168	Aaw63168	Mouse mas
175	5	55.6	3	2	AAW76946	Aaw76946	Fusion im
176	5	55.6	3	2	AAW71120	Aaw71120	Peptide A
177	5	55.6	3	2	AAW41762	Aaw41762	Colony st
178	5	55.6	3	2	AAW41282	Aaw41282	Apoptosis
179	5	55.6	3	2	AAW41286	Aaw41286	Apoptosis
180	5	55.6	3	2	AAW41278	Aaw41278	Apoptosis
181	5	55.6	3	2	ADI30304	Adi30304	Peptide r
182	5	55.6	3	2	ADI30308	Adi30308	Peptide r
183	5	55.6	3	2	ADI30311	Adi30311	Peptide r
184	5	55.6	3	2	AAY30593	Aay30593	Amino aci
185	5	55.6	3	2	AAW84193	Aaw84193	Peptide c
186	5	55.6	3	2	AAY30549	Aay30549	Amino aci
187	5	55.6	3	2	AAY13741	Aay13741	RGD pepti
188	5	55.6	3	2	AAY32872	Aay32872	Fibronect
189	5	55.6	3	2	AAY50605	Aay50605	Resin bou
190	5	55.6	3	2	AAY50604	Aay50604	Resin bou
191	5	55.6	3	2	AAY32860	Aay32860	Fibronect
192	5	55.6	3	2	AAW78373	Aaw78373	Isoelectr
193	5	55.6	3	2	AAW88697	Aaw88697	Secreted
194	5	55.6	3	3	AAB09713	Aab09713	Insulin l
195	5	55.6	3	3	AAY66966	Aay66966	T cell an
196	5	55.6	3	3	AAY66961	Aay66961	T cell an
197	5	55.6	3	3	AAY66962	Aay66962	T cell an
198	5	55.6	3	3	AAB22907	Aab22907	Cathepsin
199	5	55.6	3	3	AAY80466	Aay80466	Cell adhe
200	5	55.6	3	3	AAY80482	Aay80482	Cell adhe
201	5	55.6	3	3	AAB23991	Aab23991	Linker pe
202	5	55.6	3	3	AAY84134	Aay84134	Peptide c
203	5	55.6	3	3	AAY79609	Aay79609	Cathepsin
204	5	55.6	3	3	AAB20585	Aab20585	Targeting
205	5	55.6	3	3	AAB01578	Aab01578	Cell bind
206	5	55.6	3	4	AAM98085	Aam98085	Human pep
207	5	55.6	3	4	AAB91977	Aab91977	Fibronect
208	5	55.6	3	4	AAB91692	Aab91692	Opioid pe
209	5	55.6	3	4	AAB91597	Aab91597	Opioid pe
210	5	55.6	3	4	AAB92005	Aab92005	FMRF neur
211	5	55.6	3	4	AAB91681	Aab91681	Opioid pe
212	5	55.6	3	4	AAB91211	Aab91211	Neuropept
213	5	55.6	3	4	AAB91994	Aab91994	Fibronect
214	5	55.6	3	4	AAB92401	Aab92401	Miscellan
215	5	55.6	3	4	ABB50531	Abb50531	Human sec
216	5	55.6	3	4	AAB61825	Aab61825	Human Ig
217	5	55.6	3	5	AAU81103	Aau81103	Integrin-
218	5	55.6	3	5	AAU81101	Aau81101	Integrin-
219	5	55.6	3	5	ABG94469	Abg94469	Protease
220	5	55.6	3	5	ABB81520	Abb81520	Neural th
221	5	55.6	3	5	ABG93564	Abg93564	Human P-g
222	5	55.6	3	5	ABG93657	Abg93657	Human Mit
223	5	55.6	3	5	ABG93647	Abg93647	Human Mit
224	5	55.6	3	5	ABG93622	Abg93622	Human P-g
225	5	55.6	3	5	ABG93534	Abg93534	Human P-g
226	5	55.6	3	5	ABG93646	Abg93646	Human Mit
227	5	55.6	3	5	ABG93536	Abg93536	Human P-g
228	5	55.6	3	5	ABG93562	Abg93562	Human P-g
229	5	55.6	3	5	AAU82936	Aau82936	Fibronect
230	5	55.6	3	5	ABG60464	Abg60464	Selective
231	5	55.6	3	5	ABG60498	Abg60498	Selective
232	5	55.6	3	5	ABG60484	Abg60484	Selective
233	5	55.6	3	5	AAM47882	Aam47882	Human hae
234	5	55.6	3	5	AAU69663	Aau69663	Cell deat
235	5	55.6	3	5	AAU98509	Aau98509	Matrix me
236	5	55.6	3	5	ABG77867	Abg77867	Targettin
237	5	55.6	3	5	ABG77886	Abg77886	Targettin
238	5	55.6	3	5	ABG77833	Abg77833	Targettin
239	5	55.6	3	5	ABG77836	Abg77836	Targettin
240	5	55.6	3	5	ABG77782	Abg77782	Targettin
241	5	55.6	3	5	ABG77479	Abg77479	Targettin
242	5	55.6	3	5	ABG77483	Abg77483	Targettin
243	5	55.6	3	5	ABG77837	Abg77837	Targettin

244	5	55.6	3	5	ABG77863	Abg77863	Targettin
245	5	55.6	3	5	ABG77883	Abg77883	Targettin
246	5	55.6	3	5	ABG77889	Abg77889	Targettin
247	5	55.6	3	5	ABG77792	Abg77792	Targettin
248	5	55.6	3	5	ABG77864	Abg77864	Targettin
249	5	55.6	3	5	ABG77785	Abg77785	Targettin
250	5	55.6	3	5	ABG77807	Abg77807	Targettin
251	5	55.6	3	5	ABG77887	Abg77887	Targettin
252	5	55.6	3	5	ABG77890	Abg77890	Targettin
253	5	55.6	3	5	ABG77786	Abg77786	Targettin
254	5	55.6	3	5	ABG77779	Abg77779	Targettin
255	5	55.6	3	5	ABG77780	Abg77780	Targettin
256	5	55.6	3	5	ABG77814	Abg77814	Targettin
257	5	55.6	3	5	ABG77845	Abg77845	Targettin
258	5	55.6	3	5	ABG77486	Abg77486	Targettin
259	5	55.6	3	5	ABG77805	Abg77805	Targettin
260	5	55.6	3	5	ABG77839	Abg77839	Targettin
261	5	55.6	3	5	ABG77852	Abg77852	Targettin
262	5	55.6	3	5	ABG77783	Abg77783	Targettin
263	5	55.6	3	5	ABG77811	Abg77811	Targettin
264	5	55.6	3	5	ABG77488	Abg77488	Targettin
265	5	55.6	3	5	ABG77775	Abg77775	Targettin
266	5	55.6	3	5	ABG77808	Abg77808	Targettin
267	5	55.6	3	5	ABG77876	Abg77876	Targettin
268	5	55.6	3	5	ABG77835	Abg77835	Targettin
269	5	55.6	3	5	ABG77843	Abg77843	Targettin
270	5	55.6	3	5	AAU85775	Aau85775	Angiotens
271	5	55.6	3	6	ABU72609	Abu72609	Novel pro
272	5	55.6	3	6	ABU72668	Abu72668	Novel pro
273	5	55.6	3	6	ABU90757	Abu90757	Peptide #
274	5	55.6	3	6	ABU13829	Abu13829	Thrombosp
275	5	55.6	3	6	ABU13855	Abu13855	Thrombosp
276	5	55.6	3	6	ABU13854	Abu13854	Thrombosp
277	5	55.6	3	6	ABU60858	Abu60858	Peptide p
278	5	55.6	3	6	ABP99827	Abp99827	Human sec
279	5	55.6	3	6	ABR39486	Abr39486	AlphaS1 c
280	5	55.6	3	6	ABU11422	Abu11422	Angiotens
281	5	55.6	3	6	ABR55066	Abr55066	MMP subst
282	5	55.6	3	6	ABR55098	Abr55098	MMP subst
283	5	55.6	3	6	ABR55099	Abr55099	MMP subst
284	5	55.6	3	6	ABR55060	Abr55060	MMP subst
285	5	55.6	3	6	ABR55101	Abr55101	MMP subst
286	5	55.6	3	6	ABR55068	Abr55068	MMP subst
287	5	55.6	3	6	ABJ19424	Abj19424	Neural th
288	5	55.6	3	6	ABU06547	Abu06547	Maize Sta
289	5	55.6	3	6	ABU12339	Abu12339	Angiotens
290	5	55.6	3	6	ABR43883	Abr43883	Synovial
291	5	55.6	3	6	ABR43890	Abr43890	Synovial
292	5	55.6	3	6	ABR43889	Abr43889	Synovial
293	5	55.6	3	6	ABR43891	Abr43891	Synovial
294	5	55.6	3	6	ABO44788	Abo44788	Novel hum
295	5	55.6	3	6	ABR82577	Abr82577	Protein k
296	5	55.6	3	6	ABR82578	Abr82578	Protein k
297	5	55.6	3	6	ABR82579	Abr82579	Protein k
298	5	55.6	3	6	ABR82606	Abr82606	Protein k
299	5	55.6	3	6	ABR82580	Abr82580	Protein k
300	5	55.6	3	7	ABR44062	Abr44062	Fragment
301	5	55.6	3	7	ABO26268	Abo26268	Human pro
302	5	55.6	3	7	ADC36727	Adc36727	Mutated v
303	5	55.6	3	7	ADC18408	Adc18408	Protease
304	5	55.6	3	7	ADE37014	Ade37014	Cell bind
305	5	55.6	3	7	ADF10724	Adf10724	Human sec
306	5	55.6	3	7	ADK41651	Adk41651	Bradikini
307	5	55.6	3	7	ADW33443	Adw33443	HLA bindi
308	5	55.6	3	7	ADW37077	Adw37077	HLA bindi
309	5	55.6	3	7	ADW36789	Adw36789	HLA bindi
310	5	55.6	3	7	ADW35709	Adw35709	HLA bindi
311	5	55.6	3	7	ADW34591	Adw34591	HLA bindi
312	5	55.6	3	7	ADW36683	Adw36683	HLA bindi
313	5	55.6	3	7	ADW33358	Adw33358	HLA bindi
314	5	55.6	3	7	ADW34676	Adw34676	HLA bindi
315	5	55.6	3	7	ADW36739	Adw36739	HLA bindi
316	5	55.6	3	7	ADW34476	Adw34476	HLA bindi
317	5	55.6	3	7	ADW37278	Adw37278	HLA bindi
318	5	55.6	3	7	ADW36968	Adw36968	HLA bindi
319	5	55.6	3	8	ADI36142	Adi36142	Human ang
320	5	55.6	3	8	ADJ62029	Adj62029	Dut prote
321	5	55.6	3	8	ADM01233	Adm01233	Targeted
322	5	55.6	3	8	ADM01247	Adm01247	Targeted
323	5	55.6	3	8	ADM01248	Adm01248	Targeted
324	5	55.6	3	8	ADM01249	Adm01249	Targeted
325	5	55.6	3	8	ADM32999	Adm32999	Human imm
326	5	55.6	3	8	ADM33004	Adm33004	Human imm
327	5	55.6	3	8	ADM33000	Adm33000	Human imm
328	5	55.6	3	8	ADN17303	Adn17303	Human gal
329	5	55.6	3	8	ADP30381	Adp30381	Human sec
330	5	55.6	3	8	ADP29325	Adp29325	Human sec
331	5	55.6	3	8	ADO34674	Ado34674	Epstein B
332	5	55.6	3	8	ADQ59611	Adq59611	Peptide l
333	5	55.6	3	8	ADR04670	Adr04670	Factor XA
334	5	55.6	3	8	ADR04679	Adr04679	Factor XA
335	5	55.6	3	8	ADS85334	Ads85334	PCAV puta
336	5	55.6	3	8	ADR68276	Adr68276	Murine 26
337	5	55.6	3	8	ADS82723	Ads82723	Arabidops
338	5	55.6	3	8	ADU15638	Adu15638	MUC1 cyto
339	5	55.6	3	8	ADU15617	Adu15617	MUC1 cyto
340	5	55.6	3	8	ADU15628	Adu15628	MUC1 cyto
341	5	55.6	3	8	ADT51230	Adt51230	G protein
342	5	55.6	3	8	ADT51239	Adt51239	G protein
343	5	55.6	3	8	ADT51228	Adt51228	G protein
344	5	55.6	3	9	ADV52535	Adv52535	Gamma-sec
345	5	55.6	3	9	ADY29695	Ady29695	Antibacte
346	5	55.6	3	9	ADY29701	Ady29701	Antibacte
347	5	55.6	3	9	ADX39036	Adx39036	Lactate d
348	5	55.6	3	9	ADY64208	Ady64208	Bovine la
349	5	55.6	3	9	ADY64204	Ady64204	Bovine la
350	5	55.6	3	9	ADY64200	Ady64200	Bovine la
351	5	55.6	3	9	ADY99974	Ady99974	Rat neura
352	5	55.6	3	9	AEA21530	Aea21530	Human ant
353	5	55.6	3	9	AEA35151	Aea35151	LPS bindi
354	5	55.6	3	9	AEA39119	Aea39119	Flaviviri
355	5	55.6	3	9	AEA39121	Aea39121	Flaviviri
356	5	55.6	3	9	AEA39117	Aea39117	Flaviviri
357	5	55.6	3	9	AEA39123	Aea39123	Flaviviri
358	5	55.6	4	1	AAP10137	Aap10137	Sequence
359	5	55.6	4	1	AAP10133	Aap10133	Sequence
360	5	55.6	4	1	AAP10129	Aap10129	Sequence
361	5	55.6	4	1	AAP10138	Aap10138	Sequence
362	5	55.6	4	1	AAP10130	Aap10130	Sequence
363	5	55.6	4	1	AAP10135	Aap10135	Sequence
364	5	55.6	4	1	AAP10132	Aap10132	Sequence
365	5	55.6	4	1	AAP10134	Aap10134	Sequence
366	5	55.6	4	1	AAP10136	Aap10136	Sequence
367	5	55.6	4	1	AAP10131	Aap10131	Sequence
368	5	55.6	4	1	AAP10139	Aap10139	Sequence
369	5	55.6	4	1	AAP10365	Aap10365	Peptide s
370	5	55.6	4	1	AAP10364	Aap10364	Peptide s
371	5	55.6	4	1	AAP10362	Aap10362	Peptide s
372	5	55.6	4	1	AAP10078	Aap10078	Sequence
373	5	55.6	4	1	AAP10080	Aap10080	Sequence
374	5	55.6	4	1	AAP10092	Aap10092	Sequence
375	5	55.6	4	1	AAP10079	Aap10079	Sequence
376	5	55.6	4	1	AAP10093	Aap10093	Sequence
377	5	55.6	4	1	AAP10094	Aap10094	Sequence
378	5	55.6	4	1	AAP20247	Aap20247	Hydrolysa
379	5	55.6	4	1	AAP20258	Aap20258	Protected
380	5	55.6	4	1	AAP20062	Aap20062	Tuftisin
381	5	55.6	4	1	AAP20059	Aap20059	Enzyme su
382	5	55.6	4	1	AAP30082	Aap30082	Sequence
383	5	55.6	4	1	AAP30083	Aap30083	Sequence
384	5	55.6	4	1	AAP30085	Aap30085	Sequence
385	5	55.6	4	1	AAP30084	Aap30084	Sequence
386	5	55.6	4	1	AAP30132	Aap30132	Sequence
387	5	55.6	4	1	AAP30127	Aap30127	Sequence
388	5	55.6	4	1	AAP30130	Aap30130	Sequence
389	5	55.6	4	1	AAP30129	Aap30129	Sequence

390	5	55.6	4	1	AAP30131	Aap30131 Sequence
391	5	55.6	4	1	AAP30133	Aap30133 Sequence
392	5	55.6	4	1	AAP30128	Aap30128 Sequence
393	5	55.6	4	1	AAP30126	Aap30126 Sequence
394	5	55.6	4	1	AAP30135	Aap30135 Sequence
395	5	55.6	4	1	AAP30088	Aap30088 Mammalian
396	5	55.6	4	1	AAP30087	Aap30087 Mammalian
397	5	55.6	4	1	AAP40482	Aap40482 Substrate
398	5	55.6	4	1	AAP40562	Aap40562 Oligopept
399	5	55.6	4	1	AAP40375	Aap40375 Sequence
400	5	55.6	4	1	AAP40456	Aap40456 N-termina
401	5	55.6	4	1	AAP40461	Aap40461 N-termina
402	5	55.6	4	1	AAP40464	Aap40464 N-termina
403	5	55.6	4	1	AAP40467	Aap40467 Enkephali
404	5	55.6	4	1	AAP40354	Aap40354 Sequence
405	5	55.6	4	1	AAP50690	Aap50690 Analgesic
406	5	55.6	4	1	AAP50336	Aap50336 Analgesic
407	5	55.6	4	1	AAP50109	Aap50109 N-termina
408	5	55.6	4	1	AAP50542	Aap50542 Synthetic
409	5	55.6	4	1	AAP50015	Aap50015 Sequence
410	5	55.6	4	1	AAP50016	Aap50016 Sequence
411	5	55.6	4	1	AAP61324	Aap61324 Sequence
412	5	55.6	4	1	AAP60484	Aap60484 Peptide w
413	5	55.6	4	1	AAP60497	Aap60497 Peptide w
414	5	55.6	4	1	AAP60848	Aap60848 Hypotensi
415	5	55.6	4	1	AAP60850	Aap60850 Hypotensi
416	5	55.6	4	1	AAP60851	Aap60851 Hypotensi
417	5	55.6	4	1	AAP60854	Aap60854 Hypotensi
418	5	55.6	4	1	AAP60855	Aap60855 Hypotensi
419	5	55.6	4	1	AAP60852	Aap60852 Hypotensi
420	5	55.6	4	1	AAP61666	Aap61666 Sequence
421	5	55.6	4	1	AAP61667	Aap61667 Sequence
422	5	55.6	4	1	AAP61669	Aap61669 Sequence
423	5	55.6	4	1	AAP61358	Aap61358 Sequence
424	5	55.6	4	1	AAP61356	Aap61356 Sequence
425	5	55.6	4	1	AAP61354	Aap61354 Sequence
426	5	55.6	4	1	AAP61491	Aap61491 Analgesic
427	5	55.6	4	1	AAP61741	Aap61741 Analgesic
428	5	55.6	4	1	AAP60157	Aap60157 Human gro
429	5	55.6	4	1	AAP61325	Aap61325 Sequence
430	5	55.6	4	1	AAP71341	Aap71341 Analgesic
431	5	55.6	4	1	AAP71700	Aap71700 Analgesic
432	5	55.6	4	1	AAP71484	Aap71484 Sequence
433	5	55.6	4	1	AAP71482	Aap71482 Sequence
434	5	55.6	4	1	AAP71483	Aap71483 Sequence
435	5	55.6	4	1	AAP71580	Aap71580 Growth ho
436	5	55.6	4	1	AAP71582	Aap71582 Growth ho
437	5	55.6	4	1	AAP71579	Aap71579 Growth ho
438	5	55.6	4	1	AAP71602	Aap71602 Hypotensi
439	5	55.6	4	1	AAP71604	Aap71604 Hypotensi
440	5	55.6	4	1	AAP71607	Aap71607 Hypotensi
441	5	55.6	4	1	AAP71601	Aap71601 Hypotensi
442	5	55.6	4	1	AAP71605	Aap71605 Hypotensi
443	5	55.6	4	1	AAP71608	Aap71608 Hypotensi
444	5	55.6	4	1	AAP80873	Aap80873 Hook regi
445	5	55.6	4	1	AAP82910	Aap82910 Acetylcho
446	5	55.6	4	1	AAP80424	Aap80424 Sequence
447	5	55.6	4	1	AAP80994	Aap80994 Sequence
448	5	55.6	4	1	AAP83217	Aap83217 Sequence
449	5	55.6	4	1	AAP82678	Aap82678 Thymopoie
450	5	55.6	4	1	AAB62306	Aab62306 Factor Xa
451	5	55.6	4	1	AAP80258	Aap80258 Sequence
452	5	55.6	4	1	AAP93521	Aap93521 Amino aci
453	5	55.6	4	1	AAP93594	Aap93594 Peptide a
454	5	55.6	4	1	AAP91371	Aap91371 Peptide a
455	5	55.6	4	1	AAP93313	Aap93313 Peptide a
456	5	55.6	4	1	AAP93596	Aap93596 Peptide a
457	5	55.6	4	1	AAP93599	Aap93599 Peptide a
458	5	55.6	4	1	AAP93597	Aap93597 Peptide a
459	5	55.6	4	1	AAP93312	Aap93312 Peptide a
460	5	55.6	4	1	AAP93311	Aap93311 Peptide a
461	5	55.6	4	1	AAP93595	Aap93595 Peptide a
462	5	55.6	4	1	AAP93598	Aap93598 Peptide a
463	5	55.6	4	1	AAP91418	Aap91418 Sequence
464	5	55.6	4	1	AAP91620	Aap91620 PNS/CNS m
465	5	55.6	4	1	AAP91630	Aap91630 PNS/CNS m
466	5	55.6	4	1	AAP91610	Aap91610 Motif use
467	5	55.6	4	1	AAP91624	Aap91624 Motif use
468	5	55.6	4	1	AAP91627	Aap91627 Motif use
469	5	55.6	4	1	AAP93304	Aap93304 Anti-thro
470	5	55.6	4	1	AAP91921	Aap91921 Anti-thro
471	5	55.6	4	1	AAP92312	Aap92312 Fibronect
472	5	55.6	4	2	AAR04481	Aar04481 Human imm
473	5	55.6	4	2	AAR05705	Aar05705 Peptide h
474	5	55.6	4	2	AAR05704	Aar05704 Peptide h
475	5	55.6	4	2	AAR05702	Aar05702 Peptide h
476	5	55.6	4	2	AAR03676	Aar03676 Peptide h
477	5	55.6	4	2	AAR03677	Aar03677 Peptide h
478	5	55.6	4	2	AAR05706	Aar05706 Peptide h
479	5	55.6	4	2	AAR03675	Aar03675 Peptide h
480	5	55.6	4	2	AAR07429	Aar07429 Vasopress
481	5	55.6	4	2	AAR07428	Aar07428 Vasopress
482	5	55.6	4	2	AAR06633	Aar06633 Peptide c
483	5	55.6	4	2	AAR04609	Aar04609 Antiviral
484	5	55.6	4	2	AAR04608	Aar04608 Antiviral
485	5	55.6	4	2	AAR05320	Aar05320 Anti-coag
486	5	55.6	4	2	AAR05860	Aar05860 Anti-coag
487	5	55.6	4	2	AAR07884	Aar07884 Peptide c
488	5	55.6	4	2	AAR07103	Aar07103 Melanocyt
489	5	55.6	4	2	AAR07102	Aar07102 Melanocyt
490	5	55.6	4	2	AAR05919	Aar05919 Low toxic
491	5	55.6	4	2	AAR05916	Aar05916 Low toxic
492	5	55.6	4	2	AAR05917	Aar05917 Low toxic
493	5	55.6	4	2	AAR15424	Aar15424 Coagulati
494	5	55.6	4	2	AAR12817	Aar12817 Renin-inh
495	5	55.6	4	2	AAR14267	Aar14267 Chromogen
496	5	55.6	4	2	AAR11744	Aar11744 Cyclic pl
497	5	55.6	4	2	AAR11745	Aar11745 Cyclic pl
498	5	55.6	4	2	AAR13257	Aar13257 Cytotoxic
499	5	55.6	4	2	AAR13258	Aar13258 Cytotoxic
500	5	55.6	4	2	AAR12870	Aar12870 Chromogen
501	5	55.6	4	2	AAR15327	Aar15327 Tuftsin a
502	5	55.6	4	2	AAR14433	Aar14433 Tuftsin a
503	5	55.6	4	2	AAR14947	Aar14947 Tuftsin a
504	5	55.6	4	2	AAR10408	Aar10408 Fibrinoge
505	5	55.6	4	2	AAR11750	Aar11750 Platelet
506	5	55.6	4	2	AAR11752	Aar11752 Platelet
507	5	55.6	4	2	AAR11751	Aar11751 Platelet
508	5	55.6	4	2	AAR12407	Aar12407 Anti-thro
509	5	55.6	4	2	AAR13808	Aar13808 Factor Xa
510	5	55.6	4	2	AAR13809	Aar13809 Factor Xa
511	5	55.6	4	2	AAR13807	Aar13807 Factor Xa
512	5	55.6	4	2	AAR13810	Aar13810 Factor Xa
513	5	55.6	4	2	AAR13803	Aar13803 Factor Xa
514	5	55.6	4	2	AAR13805	Aar13805 Factor Xa
515	5	55.6	4	2	AAR13804	Aar13804 Factor Xa
516	5	55.6	4	2	AAR13812	Aar13812 Factor Xa
517	5	55.6	4	2	AAR13806	Aar13806 Factor Xa
518	5	55.6	4	2	AAR14686	Aar14686 Hyaluroni
519	5	55.6	4	2	AAR27101	Aar27101 Sequence
520	5	55.6	4	2	AAR22519	Aar22519 Peptide o
521	5	55.6	4	2	AAR22525	Aar22525 Peptide o
522	5	55.6	4	2	AAR22527	Aar22527 Peptide o
523	5	55.6	4	2	AAR22518	Aar22518 Peptide o
524	5	55.6	4	2	AAR22520	Aar22520 Peptide o
525	5	55.6	4	2	AAR22521	Aar22521 Peptide o
526	5	55.6	4	2	AAR22522	Aar22522 Peptide o
527	5	55.6	4	2	AAR22524	Aar22524 Peptide o
528	5	55.6	4	2	AAR27701	Aar27701 Peptide o
529	5	55.6	4	2	AAR22517	Aar22517 Peptide o
530	5	55.6	4	2	AAR22526	Aar22526 Peptide o
531	5	55.6	4	2	AAR24410	Aar24410 CPase B-l
532	5	55.6	4	2	AAR27340	Aar27340 Allopurin
533	5	55.6	4	2	AAR27339	Aar27339 Allopurin
534	5	55.6	4	2	AAR22267	Aar22267 Protectiv
535	5	55.6	4	2	AAR28903	Aar28903 RGase par

536	5	55.6	4	2	AAR23935	Aar23935 Lactoferr
537	5	55.6	4	2	AAR29783	Aar29783 Selective
538	5	55.6	4	2	AAR22992	Aar22992 Glucoamyl
539	5	55.6	4	2	AAR26802	Aar26802 Triglycer
540	5	55.6	4	2	AAR24214	Aar24214 Fragment
541	5	55.6	4	2	AAR24213	Aar24213 Fragment
542	5	55.6	4	2	AAR31243	Aar31243 HIV princ
543	5	55.6	4	2	AAR27138	Aar27138 Fibronect
544	5	55.6	4	2	AAR27136	Aar27136 Fibronect
545	5	55.6	4	2	AAR27135	Aar27135 Fibronect
546	5	55.6	4	2	AAR27137	Aar27137 Fibronect
547	5	55.6	4	2	AAR27134	Aar27134 Fibronect
548	5	55.6	4	2	AAR32270	Aar32270 Soybean g
549	5	55.6	4	2	AAR26423	Aar26423 Immune en
550	5	55.6	4	2	AAR26422	Aar26422 Immune en
551	5	55.6	4	2	AAR26421	Aar26421 Immune en
552	5	55.6	4	2	AAR26424	Aar26424 Immune en
553	5	55.6	4	2	AAR23726	Aar23726 ACE inhib
554	5	55.6	4	2	AAR22935	Aar22935 Neurotrop
555	5	55.6	4	2	AAR22936	Aar22936 Neurotrop
556	5	55.6	4	2	AAR22937	Aar22937 Neurotrop
557	5	55.6	4	2	AAR25311	Aar25311 Cell cont
558	5	55.6	4	2	AAR25314	Aar25314 Cell cont
559	5	55.6	4	2	AAR25315	Aar25315 Cell cont
560	5	55.6	4	2	AAR25318	Aar25318 Cell cont
561	5	55.6	4	2	AAR25313	Aar25313 Cell cont
562	5	55.6	4	2	AAR25312	Aar25312 Cell cont
563	5	55.6	4	2	AAR25316	Aar25316 Cell cont
564	5	55.6	4	2	AAR25317	Aar25317 Cell cont
565	5	55.6	4	2	AAR30046	Aar30046 Cyclic HI
566	5	55.6	4	2	AAR31311	Aar31311 Alpha-sub
567	5	55.6	4	2	AAR31312	Aar31312 Alpha-sub
568	5	55.6	4	2	AAR33016	Aar33016 Alpha-sub
569	5	55.6	4	2	AAR21665	Aar21665 Cyclic te
570	5	55.6	4	2	AAR21664	Aar21664 Cyclic te
571	5	55.6	4	2	AAR23907	Aar23907 Peptide l
572	5	55.6	4	2	AAR25241	Aar25241 PTHrP (10
573	5	55.6	4	2	AAR20131	Aar20131 SEQ ID No
574	5	55.6	4	2	AAR26810	Aar26810 Polyethyl
575	5	55.6	4	2	AAR26812	Aar26812 Polyethyl
576	5	55.6	4	2	AAR25457	Aar25457 wohl+4. 3
577	5	55.6	4	2	AAR27863	Aar27863 Antimicro
578	5	55.6	4	2	AAR23084	Aar23084 Fluorinat
579	5	55.6	4	2	AAR26806	Aar26806 Propen-am
580	5	55.6	4	2	AAR26394	Aar26394 Sequence
581	5	55.6	4	2	AAR26392	Aar26392 Sequence
582	5	55.6	4	2	AAR26398	Aar26398 Sequence
583	5	55.6	4	2	AAR21345	Aar21345 HIV prote
584	5	55.6	4	2	AAR21000	Aar21000 HIV prote
585	5	55.6	4	2	AAR20174	Aar20174 Hypoxanth
586	5	55.6	4	2	AAR25999	Aar25999 Tetrapept
587	5	55.6	4	2	AAR26678	Aar26678 HIV-PND-p
588	5	55.6	4	2	AAR36713	Aar36713 Adhesion
589	5	55.6	4	2	AAR43239	Aar43239 Deletion
590	5	55.6	4	2	AAR33662	Aar33662 Growth ho
591	5	55.6	4	2	AAR33660	Aar33660 Growth ho
592	5	55.6	4	2	AAR33664	Aar33664 Growth ho
593	5	55.6	4	2	AAR39732	Aar39732 First typ
594	5	55.6	4	2	AAR37634	Aar37634 Sequence
595	5	55.6	4	2	AAR37635	Aar37635 Sequence
596	5	55.6	4	2	AAR37630	Aar37630 Sequence
597	5	55.6	4	2	AAR37628	Aar37628 Sequence
598	5	55.6	4	2	AAR37633	Aar37633 Sequence
599	5	55.6	4	2	AAR37627	Aar37627 Sequence
600	5	55.6	4	2	AAR37632	Aar37632 Sequence
601	5	55.6	4	2	AAR37629	Aar37629 Sequence
602	5	55.6	4	2	AAR37631	Aar37631 Sequence
603	5	55.6	4	2	AAR36523	Aar36523 C-amidate
604	5	55.6	4	2	AAR32383	Aar32383 Fibrinoge
605	5	55.6	4	2	AAR32382	Aar32382 Fibrinoge
606	5	55.6	4	2	AAR32381	Aar32381 Fibrinoge
607	5	55.6	4	2	AAR32380	Aar32380 Fibrinoge
608	5	55.6	4	2	AAR32384	Aar32384 Fibrinoge
609	5	55.6	4	2	AAR32386	Aar32386 Fibrinoge
610	5	55.6	4	2	AAR32385	Aar32385 Fibrinoge
611	5	55.6	4	2	AAR32377	Aar32377 Fibrinoge
612	5	55.6	4	2	AAR36616	Aar36616 Group I s
613	5	55.6	4	2	AAR37833	Aar37833 Cell adhe
614	5	55.6	4	2	AAR38743	Aar38743 KEX2 clea
615	5	55.6	4	2	AAR38742	Aar38742 KEX2 clea
616	5	55.6	4	2	AAR32109	Aar32109 Sequence
617	5	55.6	4	2	AAR26267	Aar26267 Cysteine
618	5	55.6	4	2	AAR26268	Aar26268 Cysteine
619	5	55.6	4	2	AAR43092	Aar43092 Diuretic
620	5	55.6	4	2	AAR43091	Aar43091 Diuretic
621	5	55.6	4	2	AAR43099	Aar43099 Diuretic
622	5	55.6	4	2	AAR43100	Aar43100 Diuretic
623	5	55.6	4	2	AAR44043	Aar44043 RGD pepti
624	5	55.6	4	2	AAR41638	Aar41638 Internali
625	5	55.6	4	2	AAR41649	Aar41649 Internali
626	5	55.6	4	2	AAR41650	Aar41650 Internali
627	5	55.6	4	2	AAR41643	Aar41643 Internali
628	5	55.6	4	2	AAR41642	Aar41642 Internali
629	5	55.6	4	2	AAR41636	Aar41636 Internali
630	5	55.6	4	2	AAR41637	Aar41637 Internali
631	5	55.6	4	2	AAR32651	Aar32651 Substance
632	5	55.6	4	2	AAR32652	Aar32652 C-termina
633	5	55.6	4	2	AAR32653	Aar32653 N-termina
634	5	55.6	4	2	AAR39390	Aar39390 Factor Xa
635	5	55.6	4	2	AAR37134	Aar37134 RGD pepti
636	5	55.6	4	2	AAR37133	Aar37133 RGD pepti
637	5	55.6	4	2	AAR37137	Aar37137 RGD pepti
638	5	55.6	4	2	AAR37138	Aar37138 RGD pepti
639	5	55.6	4	2	AAR37132	Aar37132 RGD pepti
640	5	55.6	4	2	AAR37135	Aar37135 RGD pepti
641	5	55.6	4	2	AAR37130	Aar37130 RGD pepti
642	5	55.6	4	2	AAR37136	Aar37136 RGD pepti
643	5	55.6	4	2	AAR43515	Aar43515 Sm B/B' e
644	5	55.6	4	2	AAR31392	Aar31392 Cyclic pl
645	5	55.6	4	2	AAR41987	Aar41987 Polymorph
646	5	55.6	4	2	AAR34873	Aar34873 Endotheli
647	5	55.6	4	2	AAR34874	Aar34874 Endotheli
648	5	55.6	4	2	AAR69333	Aar69333 Gp IIb/II
649	5	55.6	4	2	AAR69324	Aar69324 Gp IIb/II
650	5	55.6	4	2	AAR38485	Aar38485 Human RDS
651	5	55.6	4	2	AAR33343	Aar33343 Sequence
652	5	55.6	4	2	AAR43292	Aar43292 FGF antag
653	5	55.6	4	2	AAR35464	Aar35464 Propene-a
654	5	55.6	4	2	AAR35467	Aar35467 Propene-a
655	5	55.6	4	2	AAR37588	Aar37588 Factor Xa
656	5	55.6	4	2	AAR32185	Aar32185 Proteolyt
657	5	55.6	4	2	AAR38341	Aar38341 Diuretic
658	5	55.6	4	2	AAR38342	Aar38342 Diuretic
659	5	55.6	4	2	AAY18003	Aay18003 Cell adhe
660	5	55.6	4	2	AAR32189	Aar32189 Serine pr
661	5	55.6	4	2	AAR42570	Aar42570 Peptide c
662	5	55.6	4	2	AAR34514	Aar34514 Ala-Ala-G
663	5	55.6	4	2	AAR34512	Aar34512 Arg-Tyr-H
664	5	55.6	4	2	AAR38101	Aar38101 Protease-
665	5	55.6	4	2	AAR38116	Aar38116 Protease-
666	5	55.6	4	2	AAR38132	Aar38132 Protease-
667	5	55.6	4	2	AAR38106	Aar38106 Protease-
668	5	55.6	4	2	AAR38108	Aar38108 Protease-
669	5	55.6	4	2	AAR38124	Aar38124 Protease-
670	5	55.6	4	2	AAR38137	Aar38137 Protease-
671	5	55.6	4	2	AAR38128	Aar38128 Protease-
672	5	55.6	4	2	AAR38109	Aar38109 Protease-
673	5	55.6	4	2	AAR38136	Aar38136 Protease-
674	5	55.6	4	2	AAR38126	Aar38126 Protease-
675	5	55.6	4	2	AAR38145	Aar38145 Protease-
676	5	55.6	4	2	AAR38385	Aar38385 Cyclic HI
677	5	55.6	4	2	AAR43069	Aar43069 HDV large
678	5	55.6	4	2	AAR52978	Aar52978 Tetrapept
679	5	55.6	4	2	AAR48072	Aar48072 Biologica
680	5	55.6	4	2	AAR48071	Aar48071 Biologica
681	5	55.6	4	2	AAR48079	Aar48079 Intra/Int

682	5	55.6	4	2	AAR60693	Aar60693 H. halobi
683	5	55.6	4	2	AAR48252	Aar48252 Endoplasm
684	5	55.6	4	2	AAR48281	Aar48281 Nuclear r
685	5	55.6	4	2	AAR57422	Aar57422 Rabphilin
686	5	55.6	4	2	AAR46480	Aar46480 Fragment
687	5	55.6	4	2	AAR46803	Aar46803 Phytase d
688	5	55.6	4	2	AAR63104	Aar63104 Inflammat
689	5	55.6	4	2	AAR55805	Aar55805 Interleuk
690	5	55.6	4	2	AAR44362	Aar44362 Sequence
691	5	55.6	4	2	AAR53589	Aar53589 Pig somat
692	5	55.6	4	2	AAR42561	Aar42561 Tri-/tetr
693	5	55.6	4	2	AAR42558	Aar42558 Tri-/tetr
694	5	55.6	4	2	AAR42568	Aar42568 Tri-/tetr
695	5	55.6	4	2	AAR55421	Aar55421 Opioid ol
696	5	55.6	4	2	AAR55441	Aar55441 Opioid ol
697	5	55.6	4	2	AAR55469	Aar55469 Opioid ol
698	5	55.6	4	2	AAR55502	Aar55502 Opioid ol
699	5	55.6	4	2	AAR55435	Aar55435 Opioid ol
700	5	55.6	4	2	AAR55481	Aar55481 Opioid ol
701	5	55.6	4	2	AAR55497	Aar55497 Opioid ol
702	5	55.6	4	2	AAR55503	Aar55503 Opioid ol
703	5	55.6	4	2	AAR55455	Aar55455 Opioid ol
704	5	55.6	4	2	AAR55465	Aar55465 Opioid ol
705	5	55.6	4	2	AAR55494	Aar55494 Opioid ol
706	5	55.6	4	2	AAR55507	Aar55507 Opioid ol
707	5	55.6	4	2	AAR55417	Aar55417 Opioid ol
708	5	55.6	4	2	AAR55431	Aar55431 Opioid ol
709	5	55.6	4	2	AAR55425	Aar55425 Opioid ol
710	5	55.6	4	2	AAR55498	Aar55498 Opioid ol
711	5	55.6	4	2	AAR55419	Aar55419 Opioid ol
712	5	55.6	4	2	AAR55433	Aar55433 Opioid ol
713	5	55.6	4	2	AAR55451	Aar55451 Opioid ol
714	5	55.6	4	2	AAR55477	Aar55477 Opioid ol
715	5	55.6	4	2	AAR55506	Aar55506 Opioid ol
716	5	55.6	4	2	AAR55445	Aar55445 Opioid ol
717	5	55.6	4	2	AAR55475	Aar55475 Opioid ol
718	5	55.6	4	2	AAR55449	Aar55449 Opioid ol
719	5	55.6	4	2	AAR55493	Aar55493 Opioid ol
720	5	55.6	4	2	AAR55447	Aar55447 Opioid ol
721	5	55.6	4	2	AAR55484	Aar55484 Opioid ol
722	5	55.6	4	2	AAR55415	Aar55415 Opioid ol
723	5	55.6	4	2	AAR55437	Aar55437 Opioid ol
724	5	55.6	4	2	AAR55443	Aar55443 Opioid ol
725	5	55.6	4	2	AAR55423	Aar55423 Opioid ol
726	5	55.6	4	2	AAR55439	Aar55439 Opioid ol
727	5	55.6	4	2	AAR55485	Aar55485 Opioid ol
728	5	55.6	4	2	AAR55488	Aar55488 Opioid ol
729	5	55.6	4	2	AAR55453	Aar55453 Opioid ol
730	5	55.6	4	2	AAR55471	Aar55471 Opioid ol
731	5	55.6	4	2	AAR55479	Aar55479 Opioid ol
732	5	55.6	4	2	AAR55467	Aar55467 Opioid ol
733	5	55.6	4	2	AAR55473	Aar55473 Opioid ol
734	5	55.6	4	2	AAR55489	Aar55489 Opioid ol
735	5	55.6	4	2	AAR60534	Aar60534 Peptide c
736	5	55.6	4	2	AAR60533	Aar60533 Peptide c
737	5	55.6	4	2	AAR51007	Aar51007 Human HSP
738	5	55.6	4	2	AAR44665	Aar44665 Platelet
739	5	55.6	4	2	AAR44663	Aar44663 Platelet
740	5	55.6	4	2	AAR44667	Aar44667 Platelet
741	5	55.6	4	2	AAR44662	Aar44662 Platelet
742	5	55.6	4	2	AAR67015	Aar67015 RGD pepti
743	5	55.6	4	2	AAR56656	Aar56656 Platelet
744	5	55.6	4	2	AAR56655	Aar56655 Platelet
745	5	55.6	4	2	AAR56654	Aar56654 Platelet
746	5	55.6	4	2	AAR47431	Aar47431 ACE inhib
747	5	55.6	4	2	AAR57253	Aar57253 Cyclised
748	5	55.6	4	2	AAR62155	Aar62155 Basic/aci
749	5	55.6	4	2	AAW51593	Aaw51593 Peptide a
750	5	55.6	4	2	AAW51589	Aaw51589 Peptide a
751	5	55.6	4	2	AAW51591	Aaw51591 Peptide a
752	5	55.6	4	2	AAW51595	Aaw51595 Peptide a
753	5	55.6	4	2	AAR47186	Aar47186 Angiotens
754	5	55.6	4	2	AAR57452	Aar57452 Lactoferr
755	5	55.6	4	2	AAR57449	Aar57449 Lactoferr
756	5	55.6	4	2	AAR57453	Aar57453 Lactoferr
757	5	55.6	4	2	AAR57457	Aar57457 Lactoferr
758	5	55.6	4	2	AAR65804	Aar65804 Oligopept
759	5	55.6	4	2	AAR65802	Aar65802 Oligopept
760	5	55.6	4	2	AAR56760	Aar56760 C-terminu
761	5	55.6	4	2	AAR57899	Aar57899 Integrin
762	5	55.6	4	2	AAR57900	Aar57900 Integrin
763	5	55.6	4	2	AAR57885	Aar57885 Human vit
764	5	55.6	4	2	AAR57898	Aar57898 Integrin
765	5	55.6	4	2	AAR57897	Aar57897 Integrin
766	5	55.6	4	2	AAR48599	Aar48599 Sequence
767	5	55.6	4	2	AAR60580	Aar60580 Proteolyt
768	5	55.6	4	2	AAR61350	Aar61350 Vasopress
769	5	55.6	4	2	AAR49797	Aar49797 Sequence
770	5	55.6	4	2	AAR49795	Aar49795 Sequence
771	5	55.6	4	2	AAR49800	Aar49800 Sequence
772	5	55.6	4	2	AAR49798	Aar49798 Sequence
773	5	55.6	4	2	AAR49796	Aar49796 Sequence
774	5	55.6	4	2	AAR49799	Aar49799 Sequence
775	5	55.6	4	2	AAR49794	Aar49794 Sequence
776	5	55.6	4	2	AAR49793	Aar49793 Sequence
777	5	55.6	4	2	AAR60559	Aar60559 Tissue pl
778	5	55.6	4	2	AAR64131	Aar64131 Memory-st
779	5	55.6	4	2	AAR59927	Aar59927 Alpha 1-a
780	5	55.6	4	2	AAR54863	Aar54863 Factor Xa
781	5	55.6	4	2	AAR45689	Aar45689 Selectin
782	5	55.6	4	2	AAR66128	Aar66128 Peptide d
783	5	55.6	4	2	AAR66127	Aar66127 Peptide d
784	5	55.6	4	2	AAR48522	Aar48522 Lactoferr
785	5	55.6	4	2	AAR48526	Aar48526 Lactoferr
786	5	55.6	4	2	AAR48518	Aar48518 Lactoferr
787	5	55.6	4	2	AAR48521	Aar48521 Lactoferr
788	5	55.6	4	2	AAR63967	Aar63967 C5a anaph
789	5	55.6	4	2	AAR63976	Aar63976 C5a anaph
790	5	55.6	4	2	AAR63963	Aar63963 C5a anaph
791	5	55.6	4	2	AAR63971	Aar63971 C5a anaph
792	5	55.6	4	2	AAR64001	Aar64001 C5a anaph
793	5	55.6	4	2	AAR63973	Aar63973 C5a anaph
794	5	55.6	4	2	AAR63969	Aar63969 C5a anaph
795	5	55.6	4	2	AAR63979	Aar63979 C5a anaph
796	5	55.6	4	2	AAR51936	Aar51936 Helix pom
797	5	55.6	4	2	AAR51932	Aar51932 Helix pom
798	5	55.6	4	2	AAR60592	Aar60592 Native pr
799	5	55.6	4	2	AAR60595	Aar60595 proins.KT
800	5	55.6	4	2	AAR60594	Aar60594 proins.RQ
801	5	55.6	4	2	AAR60591	Aar60591 Native pr
802	5	55.6	4	2	AAR60593	Aar60593 proins.RT
803	5	55.6	4	2	AAR61198	Aar61198 Macrophag
804	5	55.6	4	2	AAR51222	Aar51222 Platelet
805	5	55.6	4	2	AAR51221	Aar51221 Platelet
806	5	55.6	4	2	AAR51493	Aar51493 Human A-m
807	5	55.6	4	2	AAR51482	Aar51482 SV40 T an
808	5	55.6	4	2	AAR51494	Aar51494 Dorsal NL
809	5	55.6	4	2	AAR51485	Aar51485 Human c-m
810	5	55.6	4	2	AAR51496	Aar51496 Dorsal cd
811	5	55.6	4	2	AAR51491	Aar51491 Human A-m
812	5	55.6	4	2	AAR54528	Aar54528 Platelet
813	5	55.6	4	2	AAR54526	Aar54526 Platelet
814	5	55.6	4	2	AAR54527	Aar54527 Platelet
815	5	55.6	4	2	AAR54529	Aar54529 Platelet
816	5	55.6	4	2	AAR60327	Aar60327 Blood coa
817	5	55.6	4	2	AAR54662	Aar54662 Native se
818	5	55.6	4	2	AAR67593	Aar67593 Tetrapept
819	5	55.6	4	2	AAR55549	Aar55549 Tetrapept
820	5	55.6	4	2	AAR67595	Aar67595 Tetrapept
821	5	55.6	4	2	AAR67594	Aar67594 Tetrapept
822	5	55.6	4	2	AAR48940	Aar48940 Thrombin
823	5	55.6	4	2	AAR48942	Aar48942 Thrombin
824	5	55.6	4	2	AAR54579	Aar54579 Cholecyst
825	5	55.6	4	2	AAR54581	Aar54581 Cholecyst
826	5	55.6	4	2	AAR54583	Aar54583 Cholecyst
827	5	55.6	4	2	AAR54584	Aar54584 Cholecyst

828	5	55.6	4	2	AAR54585	Aar54585 Cholecyst
829	5	55.6	4	2	AAR54580	Aar54580 Cholecyst
830	5	55.6	4	2	AAR54573	Aar54573 Cholecyst
831	5	55.6	4	2	AAR54570	Aar54570 Cholecyst
832	5	55.6	4	2	AAR54582	Aar54582 Cholecyst
833	5	55.6	4	2	AAR54571	Aar54571 Cholecyst
834	5	55.6	4	2	AAR54586	Aar54586 Cholecyst
835	5	55.6	4	2	AAR54572	Aar54572 Cholecyst
836	5	55.6	4	2	AAR69469	Aar69469 Integrin-
837	5	55.6	4	2	AAR69474	Aar69474 Integrin-
838	5	55.6	4	2	AAR42585	Aar42585 Kyotorphi
839	5	55.6	4	2	AAR42586	Aar42586 Kyotorphi
840	5	55.6	4	2	AAR42587	Aar42587 Kyotorphi
841	5	55.6	4	2	AAR42588	Aar42588 Kyotorphi
842	5	55.6	4	2	AAR59394	Aar59394 Thyspleni
843	5	55.6	4	2	AAR59401	Aar59401 Thyspleni
844	5	55.6	4	2	AAR59396	Aar59396 Thyspleni
845	5	55.6	4	2	AAR59390	Aar59390 Thyspleni
846	5	55.6	4	2	AAR59402	Aar59402 Thyspleni
847	5	55.6	4	2	AAR59389	Aar59389 Thyspleni
848	5	55.6	4	2	AAR59404	Aar59404 Thyspleni
849	5	55.6	4	2	AAR59393	Aar59393 Thyspleni
850	5	55.6	4	2	AAR59391	Aar59391 Thyspleni
851	5	55.6	4	2	AAR59398	Aar59398 Thyspleni
852	5	55.6	4	2	AAR59397	Aar59397 Thyspleni
853	5	55.6	4	2	AAR59409	Aar59409 Thyspleni
854	5	55.6	4	2	AAR59395	Aar59395 Thyspleni
855	5	55.6	4	2	AAR59403	Aar59403 Thyspleni
856	5	55.6	4	2	AAR59399	Aar59399 Thyspleni
857	5	55.6	4	2	AAR59408	Aar59408 Thyspleni
858	5	55.6	4	2	AAR59392	Aar59392 Thyspleni
859	5	55.6	4	2	AAR59400	Aar59400 Thyspleni
860	5	55.6	4	2	AAR57828	Aar57828 RGD contg
861	5	55.6	4	2	AAR60506	Aar60506 Factor Xa
862	5	55.6	4	2	AAR60511	Aar60511 Factor Xa
863	5	55.6	4	2	AAR60510	Aar60510 Factor Xa
864	5	55.6	4	2	AAR60507	Aar60507 Factor Xa
865	5	55.6	4	2	AAR60509	Aar60509 Factor Xa
866	5	55.6	4	2	AAR60504	Aar60504 Factor Xa
867	5	55.6	4	2	AAR60508	Aar60508 Factor Xa
868	5	55.6	4	2	AAR60512	Aar60512 Factor Xa
869	5	55.6	4	2	AAR52965	Aar52965 Human A-m
870	5	55.6	4	2	AAR52956	Aar52956 SV40 T an
871	5	55.6	4	2	AAR52968	Aar52968 Dorsal NL
872	5	55.6	4	2	AAR52959	Aar52959 Human c-m
873	5	55.6	4	2	AAR52970	Aar52970 Dorsal cd
874	5	55.6	4	2	AAR52967	Aar52967 Human A-m
875	5	55.6	4	2	AAR55046	Aar55046 161 regio
876	5	55.6	4	2	AAR55042	Aar55042 161 regio
877	5	55.6	4	2	AAR53715	Aar53715 Tetrapept
878	5	55.6	4	2	AAR53728	Aar53728 Tetrapept
879	5	55.6	4	2	AAR53719	Aar53719 Tetrapept
880	5	55.6	4	2	AAR53712	Aar53712 Tetrapept
881	5	55.6	4	2	AAR53722	Aar53722 Tetrapept
882	5	55.6	4	2	AAR53717	Aar53717 Tetrapept
883	5	55.6	4	2	AAR53721	Aar53721 Tetrapept
884	5	55.6	4	2	AAR53724	Aar53724 Tetrapept
885	5	55.6	4	2	AAR53718	Aar53718 Tetrapept
886	5	55.6	4	2	AAR53723	Aar53723 Tetrapept
887	5	55.6	4	2	AAR53726	Aar53726 Tetrapept
888	5	55.6	4	2	AAR53729	Aar53729 Tetrapept
889	5	55.6	4	2	AAR65261	Aar65261 Peptide c
890	5	55.6	4	2	AAR89742	Aar89742 C5a C-ter
891	5	55.6	4	2	AAR89743	Aar89743 C5a C-ter
892	5	55.6	4	2	AAR89741	Aar89741 C5a C-ter
893	5	55.6	4	2	AAR89740	Aar89740 C5a C-ter
894	5	55.6	4	2	AAR89739	Aar89739 C5a C-ter
895	5	55.6	4	2	AAW13702	Aaw13702 Factor Xa
896	5	55.6	4	2	AAR85941	Aar85941 Peptidase
897	5	55.6	4	2	AAR85940	Aar85940 Peptidase
898	5	55.6	4	2	AAR81695	Aar81695 Analogue
899	5	55.6	4	2	AAR71564	Aar71564 Hepatitis
900	5	55.6	4	2	AAR90219	Aar90219 Conserved
901	5	55.6	4	2	AAR82448	Aar82448 Optional
902	5	55.6	4	2	AAR76136	Aar76136 hML(152-1
903	5	55.6	4	2	AAR88292	Aar88292 Memory en
904	5	55.6	4	2	AAR61396	Aar61396 PF4-relat
905	5	55.6	4	2	AAR88219	Aar88219 Lactoferr
906	5	55.6	4	2	AAR86902	Aar86902 Thrombin
907	5	55.6	4	2	AAR86901	Aar86901 Blood coa
908	5	55.6	4	2	AAR87659	Aar87659 His-(D)Ph
909	5	55.6	4	2	AAR87663	Aar87663 His-(D)Ph
910	5	55.6	4	2	AAR77843	Aar77843 RGDV pept
911	5	55.6	4	2	AAR66944	Aar66944 RACE pept
912	5	55.6	4	2	AAR84690	Aar84690 Bovine la
913	5	55.6	4	2	AAR84686	Aar84686 Bovine la
914	5	55.6	4	2	AAR84689	Aar84689 Bovine la
915	5	55.6	4	2	AAR84694	Aar84694 Bovine la
916	5	55.6	4	2	AAR83116	Aar83116 Factor-Xa
917	5	55.6	4	2	AAW09862	Aaw09862 Thrombin
918	5	55.6	4	2	AAW09861	Aaw09861 Thrombin
919	5	55.6	4	2	AAR68591	Aar68591 Rat NDF p
920	5	55.6	4	2	AAR80049	Aar80049 Peptidase
921	5	55.6	4	2	AAR80056	Aar80056 Peptidase
922	5	55.6	4	2	AAR83236	Aar83236 Integrin-
923	5	55.6	4	2	AAR83231	Aar83231 Integrin-
924	5	55.6	4	2	AAR87691	Aar87691 Oligopept
925	5	55.6	4	2	AAR87690	Aar87690 Oligopept
926	5	55.6	4	2	AAR87695	Aar87695 Oligopept
927	5	55.6	4	2	AAR87694	Aar87694 Oligopept
928	5	55.6	4	2	AAR87689	Aar87689 Oligopept
929	5	55.6	4	2	AAR79757	Aar79757 Anti-para
930	5	55.6	4	2	AAR80260	Aar80260 Anti-para
931	5	55.6	4	2	AAR79759	Aar79759 Anti-para
932	5	55.6	4	2	AAR79760	Aar79760 Anti-para
933	5	55.6	4	2	AAR75601	Aar75601 gp120 bin
934	5	55.6	4	2	AAR75603	Aar75603 gp120 bin
935	5	55.6	4	2	AAR69104	Aar69104 Activated
936	5	55.6	4	2	AAR85667	Aar85667 Anti-alle
937	5	55.6	4	2	AAR80463	Aar80463 Artificia
938	5	55.6	4	2	AAR80464	Aar80464 Artificia
939	5	55.6	4	2	AAR79354	Aar79354 Human con
940	5	55.6	4	2	AAR79357	Aar79357 Mutant th
941	5	55.6	4	2	AAR79363	Aar79363 Mutant th
942	5	55.6	4	2	AAR79361	Aar79361 Mutant th
943	5	55.6	4	2	AAR79360	Aar79360 Mutant th
944	5	55.6	4	2	AAR69769	Aar69769 Thrombosp
945	5	55.6	4	2	AAR79362	Aar79362 Mutant th
946	5	55.6	4	2	AAR69768	Aar69768 Thrombosp
947	5	55.6	4	2	AAR69767	Aar69767 Thrombosp
948	5	55.6	4	2	AAR79356	Aar79356 Thrombosp
949	5	55.6	4	2	AAR69766	Aar69766 Thrombosp
950	5	55.6	4	2	AAR62945	Aar62945 RGD contg
951	5	55.6	4	2	AAR62947	Aar62947 RGD contg
952	5	55.6	4	2	AAR62946	Aar62946 RGD contg
953	5	55.6	4	2	AAW08853	Aaw08853 Peptide c
954	5	55.6	4	2	AAW08855	Aaw08855 Peptide c
955	5	55.6	4	2	AAW08854	Aaw08854 Peptide c
956	5	55.6	4	2	AAR64811	Aar64811 ScFv-Lys
957	5	55.6	4	2	AAR65180	Aar65180 Dibasic a
958	5	55.6	4	2	AAR89791	Aar89791 Melanotro
959	5	55.6	4	2	AAR89788	Aar89788 Melanotro
960	5	55.6	4	2	AAR81988	Aar81988 Peptidyl
961	5	55.6	4	2	AAR84895	Aar84895 ATPTR2Ap
962	5	55.6	4	2	AAR67881	Aar67881 HIV virus
963	5	55.6	4	2	AAR67887	Aar67887 HIV virus
964	5	55.6	4	2	AAR67880	Aar67880 HIV virus
965	5	55.6	4	2	AAR89771	Aar89771 Synthetic
966	5	55.6	4	2	AAR89772	Aar89772 Synthetic
967	5	55.6	4	2	AAR82908	Aar82908 Non-RGD,
968	5	55.6	4	2	AAR70473	Aar70473 Cancer me
969	5	55.6	4	2	AAR70476	Aar70476 Cancer me
970	5	55.6	4	2	AAR72663	Aar72663 Cladospor
971	5	55.6	4	2	AAW11946	Aaw11946 Dimeric p
972	5	55.6	4	2	AAR74917	Aar74917 Urea plas
973	5	55.6	4	2	AAR62426	Aar62426 Accelerat

974	5	55.6	4	2	AAR62433	Aar62433 Accelerat
975	5	55.6	4	2	AAR67280	Aar67280 Soybean a
976	5	55.6	4	2	AAR70187	Aar70187 Arg-gingi
977	5	55.6	4	2	AAR73944	Aar73944 Endotoxin
978	5	55.6	4	2	AAR73943	Aar73943 Endotoxin
979	5	55.6	4	2	AAR73942	Aar73942 Endotoxin
980	5	55.6	4	2	AAR90898	Aar90898 Mu-opioid
981	5	55.6	4	2	AAR90899	Aar90899 Mu-opioid
982	5	55.6	4	2	AAR90887	Aar90887 Mu-opioid
983	5	55.6	4	2	AAR90893	Aar90893 Mu-opioid
984	5	55.6	4	2	AAR90906	Aar90906 Mu-opioid
985	5	55.6	4	2	AAR90912	Aar90912 Mu-opioid
986	5	55.6	4	2	AAR90903	Aar90903 Mu-opioid
987	5	55.6	4	2	AAR90907	Aar90907 Mu-opioid
988	5	55.6	4	2	AAR75530	Aar75530 Peptide f
989	5	55.6	4	2	AAW29563	Aaw29563 RGD pepti
990	5	55.6	4	2	AAR83078	Aar83078 Class I M
991	5	55.6	4	2	AAR83082	Aar83082 Class I M
992	5	55.6	4	2	AAR83083	Aar83083 Class I M
993	5	55.6	4	2	AAR83084	Aar83084 Class I M
994	5	55.6	4	2	AAR83079	Aar83079 Class I M
995	5	55.6	4	2	AAR83081	Aar83081 Class I M
996	5	55.6	4	2	AAR93665	Aar93665 HIV princ
997	5	55.6	4	2	AAR79634	Aar79634 Endocardi
998	5	55.6	4	2	AAR66089	Aar66089 Myelopoie
999	5	55.6	4	2	AAR69077	Aar69077 Leaving u
1000	5	55.6	4	9	AEB46028	Aeb46028 Amino aci

ALIGNMENTS

RESULT 1
ADH29628
ID ADH29628 standard; peptide; 2 AA.
XX
AC ADH29628;
XX
DT 11-MAR-2004 (first entry)
XX
DE Swinepox homology vector 538-46.26 junction region B peptide.
XX
KW Swinepox virus; viral vector; homology vector; vaccine; antigen; tumour;
KW cytokine; immune response; feline immunodeficiency virus infection;
KW heartworm.
XX
OS Synthetic.
OS Swinepox virus.
OS Newcastle disease virus.
XX
PN WO9622363-A1.
XX
PD 25-JUL-1996.
XX
PF 19-JAN-1996; 96WO-US001485.
XX
PR 19-JAN-1995; 95US-00375992.
PR 07-JUN-1995; 95US-00472679.
PR 07-JUN-1995; 95US-00480640.
PR 07-JUN-1995; 95US-00488237.
XX
PA (SYTR) SYNTRO CORP.
XX
PI Cochran MD, Junker DE;
XX
DR WPI; 1996-354520/35.
DR N-PSDB; ADH29627.
XX
PT Recombinant swine:pox virus contg. foreign DNA sequence - useful for
PT delivery of vaccinating antigens or other therapeutic agents to humans or
PT animals.
XX
PS Disclosure; SEQ ID NO 33; 502pp; English.
XX
CC The invention relates to a new recombinant swinepox virus (SPV)
CC comprising, inserted into a HindIII M, N or K fragment of the SPV genome,
CC a foreign DNA sequence that can be expressed in a SPV-infected host cell.
CC Also new are homology vectors for production of recombinant SPV
CC comprising double-stranded foreign sequence with, on both sides, double-
CC stranded SPV DNA homologous to the viral genome on either side of the
CC HindIII N fragment. The recombinant SPV are vectors for delivering
CC vaccinating antigens or therapeutic agents to humans, other mammals or
CC birds. The foreign dNA sequence may encode an antigen from an infectious
CC agent or tumour, or a cytokine to stimulate an immune response. SPV can
CC also be used as diagnostic reagents, e.g. to detect feline
CC immunodeficiency virus of D. immitis (heartworm) infection. SPV is only
CC weakly pathogenic, species specific and induces an immune response. The
CC present sequence is encoded by a junction region or fragment from a SPV
CC homology vector of the invention.
XX
SQ Sequence 2 AA;

  Query Match 55.6%; Score 5; DB 2; Length 2;
  Best Local Similarity 100.0%; Pred. No. 1.9e+06;
  Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
      |
Db 2 R 2

RESULT 2
AAM98426
ID AAM98426 standard; peptide; 2 AA.
XX
AC AAM98426;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human peptide #1701 encoded by a SNP oligonucleotide.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US035498.
XX
PR 28-DEC-1999; 99US-0173419P.
PR 27-DEC-2000; 2000US-00173419.
XX
PA (CURA-) CURAGEN CORP.
XX
PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g. cancer,
PT autoimmune diseases and infections.
XX
PS Disclosure; Page 4041; 4143pp; English.
XX
CC The present invention relates to oligonucleotides (see AAL26793-AAL34659)
CC encoding polymorphic variants of proteins related to amylases, amyloid
CC proteins, angiopoietin, apoptosis related proteins, cadherin, cyclin,

CC polymerase, oncogenes, histones, kinases, colony stimulating factors,
CC complement related proteins, cytochromes, kinesins, cytokines,
CC interferons, interleukins, G-protein coupled receptors and thioesterases.
CC The present sequence is a peptide encoded by one such oligonucleotide.
CC The oligonucleotides and the peptides encoded by them may be used in the
CC prevention, diagnosis and treatment of diseases associated with
CC inappropriate expression of the proteins listed above. Disorders that may
CC be prevented, diagnosed and/or treated include multifactorial diseases
CC with a genetic component, such as autoimmune diseases (e.g. rheumatoid
CC arthritis, multiple sclerosis, diabetes, systemic lupus erythromatosus
CC and Grave's disease), inflammation, cancer (e.g. cancers of the bladder,
CC brain, breast, colon and kidney, leukaemia), diseases of the nervous
CC system and an infection of pathogenic organisms
XX
SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 4; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 1 R 1

RESULT 3
AAB91738
ID AAB91738 standard; peptide; 2 AA.
XX
AC AAB91738;
XX
DT 22-JUN-2001 (first entry)
XX
DE Opioid peptide SEQ ID NO:914.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US013576.
XX
PR 17-MAY-1999; 99US-0134406P.
PR 10-SEP-1999; 99US-0153406P.
PR 15-OCT-1999; 99US-0159783P.
XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity.
XX
PS Disclosure; Page 492; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity in
CC vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention
XX
SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 4; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 4
ABG93548
ID ABG93548 standard; peptide; 2 AA.
XX
AC ABG93548;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human P-glycoprotein tryptic peptide #68.
XX
KW Human; P-glycoprotein; tryptic digest; proteolytic cleavage product;
KW diabetes; Parkinson's disease; Alzheimer's disease; malaria; cholera;
KW human immunodeficiency virus infection; influenza; rabies; diphtheria;
KW cancer; multi-drug resistance; MDR.
XX
OS Homo sapiens.
XX
PN EP1223534-A1.
XX
PD 17-JUL-2002.
XX
PF 11-JAN-2002; 2002EP-00075095.
XX
PR 14-JAN-2001; 2001IL-00140881.
PR 19-OCT-2001; 2001US-00982172.
XX
PA (KATZ/) KATZ E I.
XX
PI Katz EI;
XX
DR WPI; 2002-645691/70.
XX
PT Generating amino acid sequences representative of desired polypeptide, by
PT computationally generating proteolytic cleavage products, analyzing and
PT selecting the set of products, thus generating amino acid sequences.
XX
PS Example 1; Page 14; 124pp; English.
XX
CC The invention relates to generating set of amino acid sequences (AAS)
CC representative of one desired polypeptide (I), involving computationally
CC generating a number of proteolytic cleavage products (PCP) from (I),
CC analysing the PCP according to one parameter defining a characteristic of
CC AAS and selecting a set of PCP according to a preset criteria for each
CC parameter, thus generating the set of AAS representative of (I). Also
CC included are (1) a computer readable storage media (II) comprising a
CC database of amino acid sequences corresponding to the polypeptide of
CC interest; (2) a system (III) for generating a database of amino acid
CC sequences corresponding to a polypeptide of interest, comprises a
CC processing unit which executes a software application configured for
CC generating the number of proteolytic cleavage products from one
CC polypeptide of interest, and analysing the number of proteolytic cleavage
CC products according to one parameter defining a characteristic of amino
CC acid sequence; (3) a kit for quantifying at least one polypeptide of
CC interest, comprises a number of peptides or antibodies each capable of

CC specifically recognising at least one peptide, where the number of
CC peptides is generated according to information derived from computational
CC analysis of the polypeptide of interest; and (4) quantifying one
CC polypeptide of interest in a biological sample, involving contacting the
CC biological sample with proteolytic agent, so as to obtain a proteolysed
CC biological sample, contacting the proteolysed biological sample with at
CC least one antibody and at least one peptide of a number of peptides, and
CC detecting presence, absence and/or level of antibody binding to thus
CC quantify one polypeptide of interest in the biological sample. The method
CC is useful for generating at least one antibody specific to a polypeptide
CC of interest. The peptides or antibodies generated may be used to diagnose
CC diabetes, Parkinson's disease, Alzheimer's disease, human
CC immunodeficiency virus infection, malaria, cholera, influenza, rabies,
CC diphtheria, cancer (e.g. breast, colon, cervix, melanoma, lung, ovary,
CC pancreas, prostate, lymphomas and leukaemias). The present sequence is a
CC predicted tryptic peptide from human P-glycoprotein generated to form
CC part of a kit for identifying multi-drug (MDR) resistance associated
CC proteins
XX
SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 5; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 5
ABG93543
ID ABG93543 standard; peptide; 2 AA.
XX
AC ABG93543;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human P-glycoprotein tryptic peptide #63.
XX
KW Human; P-glycoprotein; tryptic digest; proteolytic cleavage product;
KW diabetes; Parkinson's disease; Alzheimer's disease; malaria; cholera;
KW human immunodeficiency virus infection; influenza; rabies; diphtheria;
KW cancer; multi-drug resistance; MDR.
XX
OS Homo sapiens.
XX
PN EP1223534-A1.
XX
PD 17-JUL-2002.
XX
PF 11-JAN-2002; 2002EP-00075095.
XX
PR 14-JAN-2001; 2001IL-00140881.
PR 19-OCT-2001; 2001US-00982172.
XX
PA (KATZ/) KATZ E I.
XX
PI Katz EI;
XX
DR WPI; 2002-645691/70.
XX
PT Generating amino acid sequences representative of desired polypeptide, by
PT computationally generating proteolytic cleavage products, analyzing and
PT selecting the set of products, thus generating amino acid sequences.
XX
PS Example 1; Page 14; 124pp; English.
XX
CC The invention relates to generating set of amino acid sequences (AAS)
CC representative of one desired polypeptide (I), involving computationally
CC generating a number of proteolytic cleavage products (PCP) from (I),
CC analysing the PCP according to one parameter defining a characteristic of
CC AAS and selecting a set of PCP according to a preset criteria for each
CC parameter, thus generating the set of AAS representative of (I). Also
CC included are (1) a computer readable storage media (II) comprising a
CC database of amino acid sequences corresponding to the polypeptide of
CC interest; (2) a system (III) for generating a database of amino acid
CC sequences corresponding to a polypeptide of interest, comprises a
CC processing unit which executes a software application configured for
CC generating the number of proteolytic cleavage products from one
CC polypeptide of interest, and analysing the number of proteolytic cleavage
CC products according to one parameter defining a characteristic of amino
CC acid sequence; (3) a kit for quantifying at least one polypeptide of
CC interest, comprises a number of peptides or antibodies each capable of
CC specifically recognising at least one peptide, where the number of
CC peptides is generated according to information derived from computational
CC analysis of the polypeptide of interest; and (4) quantifying one
CC polypeptide of interest in a biological sample, involving contacting the
CC biological sample with proteolytic agent, so as to obtain a proteolysed
CC biological sample, contacting the proteolysed biological sample with at
CC least one antibody and at least one peptide of a number of peptides, and
CC detecting presence, absence and/or level of antibody binding to thus
CC quantify one polypeptide of interest in the biological sample. The method
CC is useful for generating at least one antibody specific to a polypeptide
CC of interest. The peptides or antibodies generated may be used to diagnose
CC diabetes, Parkinson's disease, Alzheimer's disease, human
CC immunodeficiency virus infection, malaria, cholera, influenza, rabies,
CC diphtheria, cancer (e.g. breast, colon, cervix, melanoma, lung, ovary,
CC pancreas, prostate, lymphomas and leukaemias). The present sequence is a
CC predicted tryptic peptide from human P-glycoprotein generated to form
CC part of a kit for identifying multi-drug (MDR) resistance associated
CC proteins
XX
SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 5; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 6
ABG93498
ID ABG93498 standard; peptide; 2 AA.
XX
AC ABG93498;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human P-glycoprotein tryptic peptide #18.
XX
KW Human; P-glycoprotein; tryptic digest; proteolytic cleavage product;
KW diabetes; Parkinson's disease; Alzheimer's disease; malaria; cholera;
KW human immunodeficiency virus infection; influenza; rabies; diphtheria;
KW cancer; multi-drug resistance; MDR.
XX
OS Homo sapiens.
XX
PN EP1223534-A1.
XX
PD 17-JUL-2002.
XX
PF 11-JAN-2002; 2002EP-00075095.
XX
PR 14-JAN-2001; 2001IL-00140881.
PR 19-OCT-2001; 2001US-00982172.
XX
PA (KATZ/) KATZ E I.
XX
PI Katz EI;
XX
DR WPI; 2002-645691/70.
XX

PT Generating amino acid sequences representative of desired polypeptide, by
PT computationally generating proteolytic cleavage products, analyzing and
PT selecting the set of products, thus generating amino acid sequences.
XX
PS Example 1; Page 13; 124pp; English.
XX
CC The invention relates to generating set of amino acid sequences (AAS)
CC representative of one desired polypeptide (I), involving computationally
CC generating a number of proteolytic cleavage products (PCP) from (I),
CC analysing the PCP according to one parameter defining a characteristic of
CC AAS and selecting a set of PCP according to a preset criteria for each
CC parameter, thus generating the set of AAS representative of (I). Also
CC included are (1) a computer readable storage media (II) comprising a
CC database of amino acid sequences corresponding to the polypeptide of
CC interest; (2) a system (III) for generating a database of amino acid
CC sequences corresponding to a polypeptide of interest, comprises a
CC processing unit which executes a software application configured for
CC generating the number of proteolytic cleavage products from one
CC polypeptide of interest, and analysing the number of proteolytic cleavage
CC products according to one parameter defining a characteristic of amino
CC acid sequence; (3) a kit for quantifying at least one polypeptide of
CC interest, comprises a number of peptides or antibodies each capable of
CC specifically recognising at least one peptide, where the number of
CC peptides is generated according to information derived from computational
CC analysis of the polypeptide of interest; and (4) quantifying one
CC polypeptide of interest in a biological sample, involving contacting the
CC biological sample with proteolytic agent, so as to obtain a proteolysed
CC biological sample, contacting the proteolysed biological sample with at
CC least one antibody and at least one peptide of a number of peptides, and
CC detecting presence, absence and/or level of antibody binding to thus
CC quantify one polypeptide of interest in the biological sample. The method
CC is useful for generating at least one antibody specific to a polypeptide
CC of interest. The peptides or antibodies generated may be used to diagnose
CC diabetes, Parkinson's disease, Alzheimer's disease, human
CC immunodeficiency virus infection, malaria, cholera, influenza, rabies,
CC diphtheria, cancer (e.g. breast, colon, cervix, melanoma, lung, ovary,
CC pancreas, prostate, lymphomas and leukaemias). The present sequence is a
CC predicted tryptic peptide from human P-glycoprotein generated to form
CC part of a kit for identifying multi-drug (MDR) resistance associated
CC proteins
XX
SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 5; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 7
ABG93615
ID ABG93615 standard; peptide; 2 AA.
XX
AC ABG93615;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human P-glycoprotein tryptic peptide #135.
XX
KW Human; P-glycoprotein; tryptic digest; proteolytic cleavage product;
KW diabetes; Parkinson's disease; Alzheimer's disease; malaria; cholera;
KW human immunodeficiency virus infection; influenza; rabies; diphtheria;
KW cancer; multi-drug resistance; MDR.
XX
OS Homo sapiens.
XX
PN EP1223534-A1.
XX
PD 17-JUL-2002.
XX
PF 11-JAN-2002; 2002EP-00075095.
XX
PR 14-JAN-2001; 2001IL-00140881.
PR 19-OCT-2001; 2001US-00982172.
XX
PA (KATZ/) KATZ E I.
XX
PI Katz EI;
XX
DR WPI; 2002-645691/70.
XX
PT Generating amino acid sequences representative of desired polypeptide, by
PT computationally generating proteolytic cleavage products, analyzing and
PT selecting the set of products, thus generating amino acid sequences.
XX
PS Example 1; Page 15; 124pp; English.
XX
CC The invention relates to generating set of amino acid sequences (AAS)
CC representative of one desired polypeptide (I), involving computationally
CC generating a number of proteolytic cleavage products (PCP) from (I),
CC analysing the PCP according to one parameter defining a characteristic of
CC AAS and selecting a set of PCP according to a preset criteria for each
CC parameter, thus generating the set of AAS representative of (I). Also
CC included are (1) a computer readable storage media (II) comprising a
CC database of amino acid sequences corresponding to the polypeptide of
CC interest; (2) a system (III) for generating a database of amino acid
CC sequences corresponding to a polypeptide of interest, comprises a
CC processing unit which executes a software application configured for
CC generating the number of proteolytic cleavage products from one
CC polypeptide of interest, and analysing the number of proteolytic cleavage
CC products according to one parameter defining a characteristic of amino
CC acid sequence; (3) a kit for quantifying at least one polypeptide of
CC interest, comprises a number of peptides or antibodies each capable of
CC specifically recognising at least one peptide, where the number of
CC peptides is generated according to information derived from computational
CC analysis of the polypeptide of interest; and (4) quantifying one
CC polypeptide of interest in a biological sample, involving contacting the
CC biological sample with proteolytic agent, so as to obtain a proteolysed
CC biological sample, contacting the proteolysed biological sample with at
CC least one antibody and at least one peptide of a number of peptides, and
CC detecting presence, absence and/or level of antibody binding to thus
CC quantify one polypeptide of interest in the biological sample. The method
CC is useful for generating at least one antibody specific to a polypeptide
CC of interest. The peptides or antibodies generated may be used to diagnose
CC diabetes, Parkinson's disease, Alzheimer's disease, human
CC immunodeficiency virus infection, malaria, cholera, influenza, rabies,
CC diphtheria, cancer (e.g. breast, colon, cervix, melanoma, lung, ovary,
CC pancreas, prostate, lymphomas and leukaemias). The present sequence is a
CC predicted tryptic peptide from human P-glycoprotein generated to form
CC part of a kit for identifying multi-drug (MDR) resistance associated
CC proteins
XX
SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 5; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 8
ABG93621
ID ABG93621 standard; peptide; 2 AA.
XX
AC ABG93621;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human P-glycoprotein tryptic peptide #141.
XX

KW Human; P-glycoprotein; tryptic digest; proteolytic cleavage product;
KW diabetes; Parkinson's disease; Alzheimer's disease; malaria; cholera;
KW human immunodeficiency virus infection; influenza; rabies; diphtheria;
KW cancer; multi-drug resistance; MDR.
XX
OS Homo sapiens.
XX
PN EP1223534-A1.
XX
PD 17-JUL-2002.
XX
PF 11-JAN-2002; 2002EP-00075095.
XX
PR 14-JAN-2001; 2001IL-00140881.
PR 19-OCT-2001; 2001US-00982172.
XX
PA (KATZ/) KATZ E I.
XX
PI Katz EI;
XX
DR WPI; 2002-645691/70.
XX
PT Generating amino acid sequences representative of desired polypeptide, by
PT computationally generating proteolytic cleavage products, analyzing and
PT selecting the set of products, thus generating amino acid sequences.
XX
PS Example 1; Page 15; 124pp; English.
XX
CC The invention relates to generating set of amino acid sequences (AAS)
CC representative of one desired polypeptide (I), involving computationally
CC generating a number of proteolytic cleavage products (PCP) from (I),
CC analysing the PCP according to one parameter defining a characteristic of
CC AAS and selecting a set of PCP according to a preset criteria for each
CC parameter, thus generating the set of AAS representative of (I). Also
CC included are (1) a computer readable storage media (II) comprising a
CC database of amino acid sequences corresponding to the polypeptide of
CC interest; (2) a system (III) for generating a database of amino acid
CC sequences corresponding to a polypeptide of interest, comprises a
CC processing unit which executes a software application configured for
CC generating the number of proteolytic cleavage products from one
CC polypeptide of interest, and analysing the number of proteolytic cleavage
CC products according to one parameter defining a characteristic of amino
CC acid sequence; (3) a kit for quantifying at least one polypeptide of
CC interest, comprises a number of peptides or antibodies each capable of
CC specifically recognising at least one peptide, where the number of
CC peptides is generated according to information derived from computational
CC analysis of the polypeptide of interest; and (4) quantifying one
CC polypeptide of interest in a biological sample, involving contacting the
CC biological sample with proteolytic agent, so as to obtain a proteolysed
CC biological sample, contacting the proteolysed biological sample with at
CC least one antibody and at least one peptide of a number of peptides, and
CC detecting presence, absence and/or level of antibody binding to thus
CC quantify one polypeptide of interest in the biological sample. The method
CC is useful for generating at least one antibody specific to a polypeptide
CC of interest. The peptides or antibodies generated may be used to diagnose
CC diabetes, Parkinson's disease, Alzheimer's disease, human
CC immunodeficiency virus infection, malaria, cholera, influenza, rabies,
CC diphtheria, cancer (e.g. breast, colon, cervix, melanoma, lung, ovary,
CC pancreas, prostate, lymphomas and leukaemias). The present sequence is a
CC predicted tryptic peptide from human P-glycoprotein generated to form
CC part of a kit for identifying multi-drug (MDR) resistance associated
CC proteins
XX
SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 5; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 9
ABG93576
ID ABG93576 standard; peptide; 2 AA.
XX
AC ABG93576;
XX
DT 25-NOV-2002 (first entry)
XX
DE Human P-glycoprotein tryptic peptide #96.
XX
KW Human; P-glycoprotein; tryptic digest; proteolytic cleavage product;
KW diabetes; Parkinson's disease; Alzheimer's disease; malaria; cholera;
KW human immunodeficiency virus infection; influenza; rabies; diphtheria;
KW cancer; multi-drug resistance; MDR.
XX
OS Homo sapiens.
XX
PN EP1223534-A1.
XX
PD 17-JUL-2002.
XX
PF 11-JAN-2002; 2002EP-00075095.
XX
PR 14-JAN-2001; 2001IL-00140881.
PR 19-OCT-2001; 2001US-00982172.
XX
PA (KATZ/) KATZ E I.
XX
PI Katz EI;
XX
DR WPI; 2002-645691/70.
XX
PT Generating amino acid sequences representative of desired polypeptide, by
PT computationally generating proteolytic cleavage products, analyzing and
PT selecting the set of products, thus generating amino acid sequences.
XX
PS Example 1; Page 14; 124pp; English.
XX
CC The invention relates to generating set of amino acid sequences (AAS)
CC representative of one desired polypeptide (I), involving computationally
CC generating a number of proteolytic cleavage products (PCP) from (I),
CC analysing the PCP according to one parameter defining a characteristic of
CC AAS and selecting a set of PCP according to a preset criteria for each
CC parameter, thus generating the set of AAS representative of (I). Also
CC included are (1) a computer readable storage media (II) comprising a
CC database of amino acid sequences corresponding to the polypeptide of
CC interest; (2) a system (III) for generating a database of amino acid
CC sequences corresponding to a polypeptide of interest, comprises a
CC processing unit which executes a software application configured for
CC generating the number of proteolytic cleavage products from one
CC polypeptide of interest, and analysing the number of proteolytic cleavage
CC products according to one parameter defining a characteristic of amino
CC acid sequence; (3) a kit for quantifying at least one polypeptide of
CC interest, comprises a number of peptides or antibodies each capable of
CC specifically recognising at least one peptide, where the number of
CC peptides is generated according to information derived from computational
CC analysis of the polypeptide of interest; and (4) quantifying one
CC polypeptide of interest in a biological sample, involving contacting the
CC biological sample with proteolytic agent, so as to obtain a proteolysed
CC biological sample, contacting the proteolysed biological sample with at
CC least one antibody and at least one peptide of a number of peptides, and
CC detecting presence, absence and/or level of antibody binding to thus
CC quantify one polypeptide of interest in the biological sample. The method
CC is useful for generating at least one antibody specific to a polypeptide
CC of interest. The peptides or antibodies generated may be used to diagnose
CC diabetes, Parkinson's disease, Alzheimer's disease, human
CC immunodeficiency virus infection, malaria, cholera, influenza, rabies,
CC diphtheria, cancer (e.g. breast, colon, cervix, melanoma, lung, ovary,
CC pancreas, prostate, lymphomas and leukaemias). The present sequence is a
CC predicted tryptic peptide from human P-glycoprotein generated to form
CC part of a kit for identifying multi-drug (MDR) resistance associated
CC proteins

```
XX
SQ   Sequence 2 AA;

  Query Match             55.6%;  Score 5;  DB 5;  Length 2;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 10
ABG93550
ID   ABG93550 standard; peptide; 2 AA.
XX
AC   ABG93550;
XX
DT   25-NOV-2002  (first entry)
XX
DE   Human P-glycoprotein tryptic peptide #70.
XX
KW   Human; P-glycoprotein; tryptic digest; proteolytic cleavage product;
KW   diabetes; Parkinson's disease; Alzheimer's disease; malaria; cholera;
KW   human immunodeficiency virus infection; influenza; rabies; diphtheria;
KW   cancer; multi-drug resistance; MDR.
XX
OS   Homo sapiens.
XX
PN   EP1223534-A1.
XX
PD   17-JUL-2002.
XX
PF   11-JAN-2002; 2002EP-00075095.
XX
PR   14-JAN-2001; 2001IL-00140881.
PR   19-OCT-2001; 2001US-00982172.
XX
PA   (KATZ/) KATZ E I.
XX
PI   Katz EI;
XX
DR   WPI; 2002-645691/70.
XX
PT   Generating amino acid sequences representative of desired polypeptide, by
PT   computationally generating proteolytic cleavage products, analyzing and
PT   selecting the set of products, thus generating amino acid sequences.
XX
PS   Example 1; Page 14; 124pp; English.
XX
CC   The invention relates to generating set of amino acid sequences (AAS)
CC   representative of one desired polypeptide (I), involving computationally
CC   generating a number of proteolytic cleavage products (PCP) from (I),
CC   analysing the PCP according to one parameter defining a characteristic of
CC   AAS and selecting a set of PCP according to a preset criteria for each
CC   parameter, thus generating the set of AAS representative of (I). Also
CC   included are (1) a computer readable storage media (II) comprising a
CC   database of amino acid sequences corresponding to the polypeptide of
CC   interest; (2) a system (III) for generating a database of amino acid
CC   sequences corresponding to a polypeptide of interest, comprises a
CC   processing unit which executes a software application configured for
CC   generating the number of proteolytic cleavage products from one
CC   polypeptide of interest, and analysing the number of proteolytic cleavage
CC   products according to one parameter defining a characteristic of amino
CC   acid sequence; (3) a kit for quantifying at least one polypeptide of
CC   interest, comprises a number of peptides or antibodies each capable of
CC   specifically recognising at least one peptide, where the number of
CC   peptides is generated according to information derived from computational
CC   analysis of the polypeptide of interest; and (4) quantifying one
CC   polypeptide of interest in a biological sample, involving contacting the
CC   biological sample with proteolytic agent, so as to obtain a proteolysed
CC   biological sample, contacting the proteolysed biological sample with at
CC   least one antibody and at least one peptide of a number of peptides, and
CC   detecting presence, absence and/or level of antibody binding to thus
CC   quantify one polypeptide of interest in the biological sample. The method
CC   is useful for generating at least one antibody specific to a polypeptide
CC   of interest. The peptides or antibodies generated may be used to diagnose
CC   diabetes, Parkinson's disease, Alzheimer's disease, human
CC   immunodeficiency virus infection, malaria, cholera, influenza, rabies,
CC   diphtheria, cancer (e.g. breast, colon, cervix, melanoma, lung, ovary,
CC   pancreas, prostate, lymphomas and leukaemias). The present sequence is a
CC   predicted tryptic peptide from human P-glycoprotein generated to form
CC   part of a kit for identifying multi-drug (MDR) resistance associated
CC   proteins
XX
SQ   Sequence 2 AA;

  Query Match             55.6%;  Score 5;  DB 5;  Length 2;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 11
ABP53564
ID   ABP53564 standard; peptide; 2 AA.
XX
AC   ABP53564;
XX
DT   16-DEC-2002  (first entry)
XX
DE   Plasmid pRSET SfiI/Not I scFv/Angiogenin version 2 peptide #3.
XX
KW   Antibody; leukaemia; JL-1.
XX
OS   Enterobacteria phage T7.
OS   Synthetic.
XX
PN   KR2002005128-A.
XX
PD   17-JAN-2002.
XX
PF   08-JUL-2000; 2000KR-00039110.
XX
PR   08-JUL-2000; 2000KR-00039110.
XX
PA   (HANW-) HANWHA CHEM CORP.
XX
PI   Ju HM,  Jung JH,  Kim SE;
XX
DR   WPI; 2002-544400/58.
DR   N-PSDB; ABQ82242.
XX
PT   Producing specific antibody derivatives for JL-1, useful for diagnosing
PT   and treating leukemia.
XX
PS   Disclosure; Page 13; 14pp; Korean.
XX
CC   The present invention describes a method for producing specific antibody
CC   derivatives for JL-1, thereby diagnosing and treating leukaemia. The
CC   specific antibody derivatives for JL-1 (KCTC 0790BP) can have an isotope
CC   selected from iodine (125), technetium (99m), rhenium (188) optionally
CC   attached at the end of the amino acid sequence, or a toxic protein gene
CC   is optionally attached at the end of the amino acid sequence. The
CC   antibody derivatives for JL-1 are useful for diagnosing and treating
CC   leukaemia. The present sequence represents a partial plasmid peptide
CC   sequence from the present invention
XX
SQ   Sequence 2 AA;

  Query Match             55.6%;  Score 5;  DB 5;  Length 2;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 R 1
              |
Db          1 R 1
```

RESULT 12
ABR39485
ID ABR39485 standard; peptide; 2 AA.
XX
AC ABR39485;
XX
DT 12-JUN-2003 (first entry)
XX
DE AlphaS1 casein N-terminal fragment derived peptide #74.
XX
KW AlphaS1 casein; haematopoiesis; virucide; immunosuppressive; antilipemic;
KW immunostimulant; antidiabetic; anti-HIV; cytostatic; antibacterial;
KW antianemic.
XX
OS Synthetic.
OS Bos sp.
XX
PN WO2003018606-A2.
XX
PD 06-MAR-2003.
XX
PF 29-AUG-2002; 2002WO-IL000720.
XX
PR 30-AUG-2001; 2001US-00942121.
XX
PA (CHAY-) CHAY 13 MEDICAL RES GROUP NV.
XX
PI Sidelman Z;
XX
DR WPI; 2003-312868/30.
XX
PT New purified peptide sequences, useful for treating e.g. autoimmune
PT disorders, derived by fragmentation of an N terminus of alpha-S1 casein
PT or synthesis.
XX
PS Claim 4; Page 78; 190pp; English.
XX
CC Sequences ABR39485-509 represent synthetic peptides derived from the N-
CC terminal fragment (2-26 residues) of alphaS1 casein. The peptides
CC stimulate and enhance immune response, protect against viral infection,
CC normalizes serum cholesterol level, stimulates haematopoiesis and is non-
CC toxic. They are useful for treating a viral disease or infection, an
CC autoimmune disease, haematopoiesis; for inducing haematopoietic stem
CC cells proliferation and differentiation, megakaryocytopoiesis,
CC erythropoiesis, leukocytopoiesis, thrombocytopoiesis, plasma cell
CC proliferation, dendritic cell proliferation, macrophage proliferation;
CC for preventing thrombocytopenia, pancytopenia, granulocytopenia,
CC hyperlipidemia, cholesteremia, glucosuria, diabetes, AIDS, infection by
CC HIV, conditions associated with myeloablative doses of chemoradiotherapy
CC supported by autologous bone marrow or peripheral blood stem cell
CC transplantation (ASCT) or allogeneic bone marrow transplantation (BMT),
CC erythropoietin treatable conditions; for augmenting the effect of
CC erythropoietin or thrombopoietin; for treating thrombopoietin treatable
CC condition; for enhancing peripheral stem cell mobilization; for treating
CC haematological disease and deficiencies, hypercholesterolemia,
CC hyperglycemia, helper T-cell disorders, dendrite cell deficiencies,
CC haematopoietic stem cell disorders (including platelet, lymphocyte,
CC plasma cell and neutrophil disorders), pre-leukemic conditions, leukemic
CC conditions, immune system disorders resulting from chemotherapy or
CC radiation therapy, human immune system disorders resulting from treatment
CC of diseases of immune deficiency and bacterial infections,
CC myelodysplastic syndrome, aplastic anemia and bone marrow insufficiency;
CC for enhancing colonization of blood stem cells in a myeloablated
CC recipient
XX
SQ Sequence 2 AA;

  Query Match   55.6%; Score 5; DB 6; Length 2;
  Best Local Similarity 100.0%; Pred. No. 1.9e+06;
  Matches   1; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 R 1
              |
Db          1 R 1
```

RESULT 13
AAE39209
ID AAE39209 standard; peptide; 2 AA.
XX
AC AAE39209;
XX
DT 18-DEC-2003 (first entry)
XX
DE Angiogenesis inhibitor peptide motif #38.
XX
KW Angiogenesis inhibitor; methionine aminopeptidase-2; MetAP-2; trachoma;
KW autoimmune disease; rheumatoid arthritis; retinopathy of prematurity;
KW ocular disease; diabetic retinopathy; corneal graft rejection; cancer;
KW retrolental fibroplasia; neovascular glaucoma; rubeosis; ocular tumour;
KW immunosuppressive; cytostatic; ophthalmological.
XX
OS Unidentified.
XX
FH Key   Location/Qualifiers
FT Modified-site   2
FT   /note= "Arg(3-amino 3-pyriylpropionic acid)"
XX
PN US2002193298-A1.
XX
PD 19-DEC-2002.
XX
PF 05-OCT-2001; 2001US-00972772.
XX
PR 01-NOV-2000; 2000US-00704251.
XX
PA (PRAE-) PRAECIS PHARM INC.
XX
PI Olson GL, Self C, Lee L, Cook CM, Birktoft J;
XX
DR WPI; 2003-755034/71.
XX
PT New angiogenesis inhibitor compound useful for treating angiogenic
PT diseases such as cancer, diabetic retinopathy, hypoxia, ocular tumors,
PT trachoma.
XX
PS Disclosure; Page 5; 38pp; English.
XX
CC The present invention provides novel angiogenesis inhibitor compounds
CC comprising a methionine aminopeptidase (MetAP)-2 inhibitory core coupled
CC to a peptide. The invention is useful for treating an angiogenic disease
CC such as autoimmune disease such as rheumatoid arthritis and cancer. The
CC invention is also useful for treating angiogenic diseases such as ocular
CC disease e.g. diabetic retinopathy, retinopathy of prematurity, corneal
CC graft rejection, retrolental fibroplasia, neovascular glaucoma, rubeosis,
CC angiogenesis in eye associated with infection or surgical intervention,
CC ocular tumours and trachoma. The present sequence is angiogenesis
CC inhibitor peptide motif
XX
SQ Sequence 2 AA;

  Query Match   55.6%; Score 5; DB 7; Length 2;
  Best Local Similarity 100.0%; Pred. No. 1.9e+06;
  Matches   1; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

```
Qy          1 R 1
              |
Db          2 R 2
```

```
RESULT 14
ADL98393
ID   ADL98393 standard; peptide; 2 AA.
XX
AC   ADL98393;
XX
DT   20-MAY-2004  (first entry)
XX
DE   Human leukocyte antigen-B15 allotype B*1508/B*1501 peptide #27.
XX
KW   soluble human leukocyte antigen; HLA; locus-specific primer; truncation;
KW   multimeric HLA complex; bioreactor;
KW   major histocompatibility complex molecule; MHC; vaccine; HLA-B15;
KW   allotype; B*1508; B*1501.
XX
OS   Homo sapiens.
XX
PN   US2003166057-A1.
XX
PD   04-SEP-2003.
XX
PF   18-DEC-2001; 2001US-00022066.
XX
PR   17-DEC-1999;   99US-00465321.
PR   18-DEC-2000; 2000US-0256409P.
PR   18-DEC-2000; 2000US-0256410P.
PR   24-MAY-2001; 2001US-0293261P.
PR   09-OCT-2001; 2001US-0327907P.
PR   10-OCT-2001; 2001US-00974366.
XX
PA   (HILD/) HILDEBRAND W H.
PA   (PRIL/) PRILLIMAN K R.
XX
PI   Hildebrand WH,  Prilliman KR;
XX
DR   WPI; 2003-863700/80.
XX
PT   Producing soluble human leukocyte antigen molecules, for testing the
PT   functionality of peptide ligands, comprises utilizing a locus-specific
PT   primer having a stop codon incorporated into a 3' primer, or that
PT   truncates the allelic cDNA.
XX
PS   Disclosure; Fig 26; 148pp; English.
XX
CC   The invention describes a method of producing soluble human leukocyte
CC   antigen (HLA) molecules comprising utilising a locus-specific primer
CC   having a stop codon incorporated into a 3' primer, or a locus-specific
CC   primer that truncates the allelic cDNA resulting in a truncated PCR
CC   product having the coding regions encoding cytoplasmic and transmembrane
CC   domains of the allelic cDNA removed so that the truncated PCR product has
CC   a coding region encoding a soluble HLA molecule. Also described is a
CC   multimeric HLA complex comprising a substrate, and at least two soluble
CC   HLA molecules attached to the substrate and an apparatus or a bioreactor
CC   unit for producing major histocompatibility complex molecules. The
CC   methods are useful for producing soluble human leukocyte antigen (HLA)
CC   molecules. The multimeric HLA complex is useful for testing the
CC   functionality of peptide ligands bound to the soluble HLA molecules. The
CC   HLA molecules are also useful in vaccine development. This is the amino
CC   acid sequence of a human leukocyte antigen (HLA)-B15 peptide found in
CC   allotypes B*1508 and B*1501.
XX
SQ   Sequence 2 AA;

  Query Match             55.6%;  Score 5;  DB 7;  Length 2;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 15
ADW36641
ID   ADW36641 standard; peptide; 2 AA.
XX
AC   ADW36641;
XX
DT   10-MAR-2005  (first entry)
XX
DE   HLA binding epitope #7391.
XX
KW   Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;
KW   MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;
KW   viral disease; cancer.
XX
OS   Unidentified.
XX
PN   WO2003040165-A2.
XX
PD   15-MAY-2003.
XX
PF   18-OCT-2001; 2001WO-US051650.
XX
PR   19-OCT-2000; 2000US-0242350P.
PR   20-APR-2001; 2001US-0285624P.
XX
PA   (EPIM-) EPIMMUNE INC.
XX
PI   Sette A,  Sidney J,  Southwood S;
XX
DR   WPI; 2003-441519/41.
XX
PT   New composition comprising at least one peptide having allele-specific
PT   binding motifs for HLA, useful for preventing, treating or diagnosing
PT   viral diseases and cancer.
XX
PS   Claim 1; Page 52-379; 382pp; English.
XX
CC   The invention relates to a composition comprising at least one peptide
CC   having an isolated, prepared epitope selected from any of the sequences
CC   from 30 lists given in the specification. Also disclosed is a method for
CC   inducing a cytotoxic T cell response against a pre-selected antigen in a
CC   patient expressing a specific MHC class I allele by contacting cytotoxic
CC   T cells from the patient with the composition cited above. The
CC   composition comprises an epitope that is joined by an amino acid linker.
CC   The epitope is admixed or joined to a CTL or HTL epitope. The epitope is
CC   bound to an HLA molecule on the antigen-presenting cell, where when an A2
CC   -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL
CC   binds to a complex of the HLA molecule and the epitope. Specifically
CC   claimed are peptides having allele-specific binding motifs for HLA. The
CC   compositions and methods are useful for preventing, treating or
CC   diagnosing viral diseases and cancer. The peptide epitopes are useful as
CC   diagnostic agents for evaluating immune responses, for making antibodies
CC   and for evaluating efficacy of a vaccine. Sequences given in ADW29251-
CC   ADW37745 represent epitopes of the invention as given in Tables 2-31.
XX
SQ   Sequence 2 AA;

  Query Match             55.6%;  Score 5;  DB 7;  Length 2;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 16
ADW36964
ID   ADW36964 standard; peptide; 2 AA.
XX
AC   ADW36964;
XX
DT   10-MAR-2005  (first entry)
XX
```

DE HLA binding epitope #7714.
XX
KW Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;
KW MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;
KW viral disease; cancer.
XX
OS Unidentified.
XX
PN WO2003040165-A2.
XX
PD 15-MAY-2003.
XX
PF 18-OCT-2001; 2001WO-US051650.
XX
PR 19-OCT-2000; 2000US-0242350P.
PR 20-APR-2001; 2001US-0285624P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S;
XX
DR WPI; 2003-441519/41.
XX
PT New composition comprising at least one peptide having allele-specific
PT binding motifs for HLA, useful for preventing, treating or diagnosing
PT viral diseases and cancer.
XX
PS Claim 1; Page 52-379; 382pp; English.
XX
CC The invention relates to a composition comprising at least one peptide
CC having an isolated, prepared epitope selected from any of the sequences
CC from 30 lists given in the specification. Also disclosed is a method for
CC inducing a cytotoxic T cell response against a pre-selected antigen in a
CC patient expressing a specific MHC class I allele by contacting cytotoxic
CC T cells from the patient with the composition cited above. The
CC composition comprises an epitope that is joined by an amino acid linker.
CC The epitope is admixed or joined to a CTL or HTL epitope. The epitope is
CC bound to an HLA molecule on the antigen-presenting cell, where when an A2
CC -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL
CC binds to a complex of the HLA molecule and the epitope. Specifically
CC claimed are peptides having allele-specific binding motifs for HLA. The
CC compositions and methods are useful for preventing, treating or
CC diagnosing viral diseases and cancer. The peptide epitopes are useful as
CC diagnostic agents for evaluating immune responses, for making antibodies
CC and for evaluating efficacy of a vaccine. Sequences given in ADW29251-
CC ADW37745 represent epitopes of the invention as given in Tables 2-31.
XX
SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 7; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 1 R 1

RESULT 17
ADW37285
ID ADW37285 standard; peptide; 2 AA.
XX
AC ADW37285;
XX
DT 10-MAR-2005 (first entry)
XX
DE HLA binding epitope #8035.
XX
KW Virucide; cytostatic; gene therapy; vaccine; epitope; cytotoxic T cell;
KW MHC class I; CTL; HTL; A2-restricted cytotoxic lymphocyte; HLA;
KW viral disease; cancer.
XX
OS Unidentified.
XX
PN WO2003040165-A2.
XX
PD 15-MAY-2003.
XX
PF 18-OCT-2001; 2001WO-US051650.
XX
PR 19-OCT-2000; 2000US-0242350P.
PR 20-APR-2001; 2001US-0285624P.
XX
PA (EPIM-) EPIMMUNE INC.
XX
PI Sette A, Sidney J, Southwood S;
XX
DR WPI; 2003-441519/41.
XX
PT New composition comprising at least one peptide having allele-specific
PT binding motifs for HLA, useful for preventing, treating or diagnosing
PT viral diseases and cancer.
XX
PS Claim 1; Page 52-379; 382pp; English.
XX
CC The invention relates to a composition comprising at least one peptide
CC having an isolated, prepared epitope selected from any of the sequences
CC from 30 lists given in the specification. Also disclosed is a method for
CC inducing a cytotoxic T cell response against a pre-selected antigen in a
CC patient expressing a specific MHC class I allele by contacting cytotoxic
CC T cells from the patient with the composition cited above. The
CC composition comprises an epitope that is joined by an amino acid linker.
CC The epitope is admixed or joined to a CTL or HTL epitope. The epitope is
CC bound to an HLA molecule on the antigen-presenting cell, where when an A2
CC -restricted cytotoxic lymphocyte (CTL) is present, a receptor of the CTL
CC binds to a complex of the HLA molecule and the epitope. Specifically
CC claimed are peptides having allele-specific binding motifs for HLA. The
CC compositions and methods are useful for preventing, treating or
CC diagnosing viral diseases and cancer. The peptide epitopes are useful as
CC diagnostic agents for evaluating immune responses, for making antibodies
CC and for evaluating efficacy of a vaccine. Sequences given in ADW29251-
CC ADW37745 represent epitopes of the invention as given in Tables 2-31.
XX
SQ Sequence 2 AA;

Query Match 55.6%; Score 5; DB 7; Length 2;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 18
AAP30601
ID AAP30601 standard; protein; 3 AA.
XX
AC AAP30601;
XX
DT 25-MAR-2003 (revised)
DT 31-MAY-1992 (first entry)
XX
DE Sequence of beta-lactamase encoded on pULB1523.
XX
KW Emphysema therapy; proteolytic enzyme; lung disease.
XX
OS Homo sapiens.
XX
PN BE895961-A.
XX
PD 16-JUN-1983.
XX
PF 21-FEB-1983; 83BE-00210157.
XX
PR 21-FEB-1983; 83BE-00210157.

```
PR   21-FEB-1983;   83BE-00895961.
XX
PA   (REGI-) REGION WALLONNE.
PA   (EJEC-) EJECUTIVO REG VALON.
XX
DR   WPI; 1983-700089/27.
DR   N-PSDB; AAN30207.
XX
PT   Bacterial clone producing alpha-1-antitrypsin - transformed with vector
PT   contg. antitrypsin DNA.
XX
PS   Disclosure; Fig 2; 23pp; French.
XX
CC   The inventors claim double-stranded cDNA encoding AT and E.coli clones
CC   contg. AT-DNA. The cDNA is derived from mRNA extracted from human liver.
CC   Recombinant alpha-1-AT can be used for the treatment of pollutant-induced
CC   lung damage (esp. emphysema), particularly when caused by excessive
CC   release of proteolytic enzymes. Large amts. of AT can be made by
CC   cultivating the transformed bacteria. (Updated on 25-MAR-2003 to correct
CC   PR field.)
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 1;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 19
AAP90668
ID   AAP90668 standard; protein; 3 AA.
XX
AC   AAP90668;
XX
DT   10-MAR-2003  (revised)
DT   26-MAY-1990  (first entry)
XX
DE   New antihypertensive peptide.
XX
KW   Antihypertensive peptide; angiotensin-converting enzyme; ACE.
XX
OS   Synthetic.
XX
PN   JP01083096-A.
XX
PD   28-MAR-1989.
XX
PF   25-SEP-1987;   87JP-00241646.
XX
PR   25-SEP-1987;   87JP-00241646.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
DR   WPI; 1989-136272/18.
XX
PT   New antihypertensive peptide(s) - used to inhibit angiotensin-converting
PT   enzyme.
XX
PS   Disclosure; Page; 20pp; Japanese.
XX
CC   The peptide and its salts inhibit angiotensin-converting enzyme (ACE) and
CC   are useful as antihypertensives. They may be administered orally,
CC   parenterally or rectally in the form of tablets, capsules, granules,
CC   powder, syrup, suspension, suppositories, ointment, cream, gel, plaster,
CC   inhalation compsn. or injection at a dose of 0.001-1000, pref. 0.01-10,
CC   mg, 1-3 times per day. (Updated on 10-MAR-2003 to add missing OS field.)
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 1;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 20
AAP90665
ID   AAP90665 standard; protein; 3 AA.
XX
AC   AAP90665;
XX
DT   10-MAR-2003  (revised)
DT   26-MAY-1990  (first entry)
XX
DE   New antihypertensive peptide.
XX
KW   Antihypertensive peptide; angiotensin-converting enzyme; ACE.
XX
OS   Synthetic.
XX
PN   JP01083096-A.
XX
PD   28-MAR-1989.
XX
PF   25-SEP-1987;   87JP-00241646.
XX
PR   25-SEP-1987;   87JP-00241646.
XX
PA   (AJIN ) AJINOMOTO KK.
XX
DR   WPI; 1989-136272/18.
XX
PT   New antihypertensive peptide(s) - used to inhibit angiotensin-converting
PT   enzyme.
XX
PS   Disclosure; Page; 20pp; Japanese.
XX
CC   The peptide and its salts inhibit angiotensin-converting enzyme (ACE) and
CC   are useful as antihypertensives. They may be administered orally,
CC   parenterally or rectally in the form of tablets, capsules, granules,
CC   powder, syrup, suspension, suppositories, ointment, cream, gel, plaster,
CC   inhalation compsn. or injection at a dose of 0.001-1000, pref. 0.01-10,
CC   mg, 1-3 times per day. (Updated on 10-MAR-2003 to add missing OS field.)
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 1;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 21
AAP97811
ID   AAP97811 standard; protein; 3 AA.
XX
AC   AAP97811;
XX
DT   29-JUL-1992  (first entry)
XX
DE   Sequence of fragment 24, the tryptic fragment of recombinant penicillin
DE   acyltransferase (PAT) polypeptide 2.
XX
KW   Penicillin biosynthesis; enzyme; antibiotic.
XX
OS   Penicillium chrysogenum.
```

XX
PN EP336446-A.
XX
PD 11-OCT-1989.
XX
PF 07-APR-1989; 89EP-00106214.
XX
PR 08-APR-1988; 88AT-00000922.
PR 13-JUL-1988; 88AT-00001806.
PR 08-SEP-1988; 88AT-00002201.
XX
PA (BIOC) BIOCHEMIE GMBH.
XX
PI Knauseder F, Leitner E, Palma N, Weber G;
XX
DR WPI; 1989-294357/41.
XX
PT Recombinant penicillin acyl-transferase - and DNA coding for it.
XX
PS Claim 9; Page 48; 52pp; English.
XX
CC The inventors claim recombinant penicillin acyltransferase (PAT) and DNA
CC coding for PAT. PAT catalyses the last step in the biosynthesis of
CC penicillin G and penicillin V. More specifically, the coding strand of
CC the DNA has the nucleotide sequence shown below. This includes three
CC introns and codes for a PAT protein with mol. wt. ca. 40kD. Plasmid
CC vectors pBC2001 and pBC2002 are specifically claimed
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 1; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 22
AAR00718
ID AAR00718 standard; peptide; 3 AA.
XX
AC AAR00718;
XX
DT 09-JAN-2003 (revised)
DT 29-MAY-1990 (first entry)
XX
DE Core repeat of cell-adhesive protein.
XX
KW Cell adhesion; anti-metastatic agent; immunomodulation; core repeat.
XX
OS Homo sapiens.
XX
PN EP347931-A.
XX
PD 27-DEC-1989.
XX
PF 23-JUN-1989; 89EP-00111468.
XX
PR 24-JUN-1988; 88JP-00156133.
XX
PA (AZUM/) AZUMA I.
XX
PI Saiki I, Nishi N, Azuma I, Tokura S;
XX
DR WPI; 1990-001305/01.
XX
PT Polypeptide with repeated sequences of cell-adhesive protein - used as
PT anti-metastatic agent for cancer and agonist or antagonist of cell-
PT adhesive proteins.
XX
PS Claim 2; Page 14; 16pp; English.
XX
CC Peptide core is repeated 2-20 times to form a cell-adhesive protein of
CC mol. wt. 1,500-5,000. The protein is an (ant)agonist of cell-adhesive
CC proteins such as fibronectin. It has high antimetastatic activity against
CC cancer and can be used in immunomodulation, wound healing, platelet
CC aggregation inhibition and alleviati9on of neuro-disorders. See also
CC AAR00722. (Updated on 09-JAN-2003 to add missing OS field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 1 R 1

RESULT 23
AAR04607
ID AAR04607 standard; peptide; 3 AA.
XX
AC AAR04607;
XX
DT 25-MAR-2003 (revised)
DT 05-SEP-1990 (first entry)
XX
DE Antiviral agent.
XX
KW Antiviral; M2; poliovirus; polio; heatitis.
XX
OS Synthetic.
XX
PN JP02078631-A.
XX
PD 19-MAR-1990.
XX
PF 14-SEP-1988; 88JP-00228843.
XX
PR 14-SEP-1988; 88JP-00228843.
XX
PA (NIHA) NIPPON MINING CO.
XX
DR WPI; 1990-129060/17.
XX
PT Antiviral agent contg. tri:peptide (unit) - of basic aminoacid, then
PT alanine, glycine or sarcosine, and acidic aminoacid, effective against
PT virus with protein-terminated DNA or RNA.
XX
PS Disclosure; Page ?; 4pp; Japanese.
XX
CC Peptide is effective against inhibiting propogation of DNA or RNA bonded,
CC protein containing viruses eg. Poliovirus, Hepatitis virus. (Updated on
CC 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 1 R 1

RESULT 24
AAR10543
ID AAR10543 standard; protein; 3 AA.
XX
AC AAR10543;
XX

DT 25-MAR-2003 (revised)
DT 15-MAR-1991 (first entry)
XX
DE Hypotensive oligopeptide.
XX
KW Oral hypotensive; fig.
XX
OS Ficus carica.
XX
PN JP02282394-A.
XX
PD 19-NOV-1990.
XX
PF 24-APR-1989; 89JP-00104265.
XX
PR 24-APR-1989; 89JP-00104265.
XX
PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.
PA (SHOS) SHOWA SANGYO CO.
XX
DR WPI; 1991-004480/01.
XX
PT New oligopeptide hypotensive drug - based on alanine, valine, asparagine,
PT proline, isoleucine and arginine.
XX
PS Claim 1; Page 843; 9pp; Japanese.
XX
CC Peptides may be derived from extract of fig, and are hypotensive agents.
CC (Updated on 25-MAR-2003 to correct PA field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 25
AAR32271
ID AAR32271 standard; peptide; 3 AA.
XX
AC AAR32271;
XX
DT 01-APR-1993 (first entry)
XX
DE Soybean glycinin derived hypotensive #13.
XX
KW soybean; hypotensive; hypertension; high blood pressure.
XX
OS Glycine max.
XX
PN JP04297493-A.
XX
PD 21-OCT-1992.
XX
PF 13-FEB-1991; 91JP-00105245.
XX
PR 13-FEB-1991; 91JP-00105245.
XX
PA (AJIN) AJINOMOTO KK.
XX
DR WPI; 1992-401807/49.
XX
PT Novel peptide(s) derived from soybean glycinin - are useful as
PT hypotensive agents.
XX
PS Claim 1; Page 1; 5pp; Japanese.
XX
CC This peptide is one of 17 derived from soybean glycinin. It is useful as
CC a hypotensive agent in pharmaceuticals or food. It can be administered
CC orally or parenterally at a daily dose of 0.001mg to 5g. It may be
CC prepared by solid phase synthesis
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 26
AAR36707
ID AAR36707 standard; peptide; 3 AA.
XX
AC AAR36707;
XX
DT 25-MAR-2003 (revised)
DT 26-AUG-1993 (first entry)
XX
DE Adhesion formation prevention RGD-contg. peptide.
XX
KW Tissue repair; peritoneum; surgery; post-surgically; inhibition;
KW platelet aggregation; cardiovascular; orthopedic; thoracic; ophthalmic;
KW CNS; use.
XX
OS Synthetic.
XX
PN WO9308818-A1.
XX
PD 13-MAY-1993.
XX
PF 06-NOV-1992; 92WO-US009494.
XX
PR 07-NOV-1991; 91US-00789231.
XX
PA (UYSC-) UNIV SOUTHERN CALIFORNIA.
XX
PI Dizerega GS, Rodgers KE;
XX
DR WPI; 1993-167381/20.
XX
PT Prevention of adhesion formation, partic. post-surgically - comprises
PT administering a RGD-contg. peptide for a time sufficient to permit tissue
PT repair.
XX
PS Example; Page 18; 22pp; English.
XX
CC The sequence is that of an RGD-contg. peptide which is used in a method
CC for prevention of adhesion formation for a time sufficient to permit
CC tissue repair. The method is used for minimising or preventing adhesion
CC formation, partic. in the peritoneum following surgery, but also for e.g.
CC cardiovascular, orthopedic, thoracic, ophthalmic, CNS and other uses. In
CC addn., the peptide inhibits platelet aggregation and does not induce
CC inflammation or trauma at the site of administration. (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 1 R 1

RESULT 27

```
AAR53144
ID   AAR53144 standard; peptide; 3 AA.
XX
AC   AAR53144;
XX
DT   02-JUN-1994  (first entry)
XX
DE   RGD peptide derivative #6.
XX
KW   Drug; organ transplantation; rejection; immune disorder; systemic lupus.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "C9H19C0-Arg, C13H27C0-Arg, C15H31C0-Arg, CH3-
FT                   [(CH(CH3)-(CH2)3]3-CH(CH3)-CH2-CO-Arg or CH3-[(CH(CH3)-
FT                   (CH2)3]2-CH(CH3)-CH2-CO-Arg"
FT   Modified-site   4
FT                   /note= "Asp-OH or Asp-NH2"
XX
PN   JP05255105-A.
XX
PD   05-OCT-1993.
XX
PF   16-MAR-1992;   92JP-00058460.
XX
PR   16-MAR-1992;   92JP-00058460.
XX
PA   (FUJF ) FUJI PHOTO FILM CO LTD.
XX
DR   WPI; 1993-348360/44.
XX
PT   Immuno-control drug for organ transplant rejection etc. - contains
PT   peptide having arginine, glycine, aspartic acid sequence.
XX
PS   Disclosure; Page 3; 11pp; Japanese.
XX
CC   The sequences given in AAR44043-47 and AAR53144 represent examples of the
CC   claimed RGD containing peptide of the invention. These peptides all
CC   correspond to the generic formulae: HO2-(CH2)m-C(O)-([X]-Arg-Gly-Asp-
CC   [Y])n-O-CH2CH(OR1)CH2OR2 or R3-([X]-Arg-Gly-Asp-[Y])n-Z [X], [Y] = amino
CC   acid or peptide residues; m = 1-5; n = 1-5; R1, R2 = H or 8-24C acyk or
CC   alkyl; R3 = 6-24C acyl; Z = hydroxyl or amino. These peptides form the
CC   active part of drugs which are used for the control of organ
CC   transplantation rejection or immune disorders such as systemic lupus
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 28
AAR38487
ID   AAR38487 standard; protein; 3 AA.
XX
AC   AAR38487;
XX
DT   25-MAR-2003  (revised)
DT   11-NOV-1993  (first entry)
XX
DE   Human RDS Leu185Pro fragment.
XX
KW   Human; mutant; retinal degeneration; primer; rds; RDS; photoreceptor;
KW   hereditary; retinitis pigmentosa; amplify; retina; murine; autosomal;
KW   dominant; PCR.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 3
FT                   /note= "Leu185Pro"
XX
PN   WO9312134-A1.
XX
PD   24-JUN-1993.
XX
PF   08-DEC-1992;   92WO-US010536.
XX
PR   11-DEC-1991;   91US-00805123.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Dryja TP,  Berson EL;
XX
DR   WPI; 1993-214088/26.
DR   N-PSDB; AAQ43556.
XX
PT   Probe or primer contg. sequence of human retinal degeneration slow
PT   protein mutant - used to diagnose hereditary retinal degenerative
PT   diseases.
XX
PS   Disclosure; Fig 13C; 56pp; English.
XX
CC   The sequences given in AAR38485-87 represent regions of the human retinal
CC   degenaration slow (RDS) protein which contain mutations. These mutations
CC   of the human RDS protein cosegregate with autosomal dominant retinitis
CC   pigmentosa. The RDS gene sequence was isolated via the murine rds gene
CC   and has been mapped to chromosome 6p. The murine rds gene is a
CC   semidominant mutation with a phenotype of abnormal development of rod and
CC   cone photoreceptors, followed by their slow degeneration. The DNA
CC   encoding the mutation containing regions was isolated by polymerase chain
CC   reaction (PCR) using primers derived from RDS gene intron flanking
CC   sequences. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 29
AAR42569
ID   AAR42569 standard; peptide; 3 AA.
XX
AC   AAR42569;
XX
DT   25-MAR-2003  (revised)
DT   22-JUN-1994  (first entry)
XX
DE   Peptide corresponding to pseudo-substrate region of zeta-PKC.
XX
KW   Zeta-protein kinase C inhibitor; zeta-PKC; pseudosubstrate; tumour;
KW   hyperproliferative disorders; psoriasis; viral infection; HIV.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "can be N-acetylated"
XX
PN   WO9320101-A1.
XX
PD   14-OCT-1993.
XX
```

PF 02-APR-1993; 93WO-EP000816.
XX
PR 06-APR-1992; 92EP-00500034.
XX
PA (GLAX) GLAXO SA.
XX
PI Diaz-Meco Conde MT, Moscat Guillen J;
XX
DR WPI; 1993-336831/42.
XX
PT Peptide(s) corresp. to the pseudo-substrate region of zeta-PKC - used for
PT treatment of tumours, hyper-proliferative disorders and viral infections.
XX
PS Claims 4 + 5; Page 43; 57pp; English.
XX
CC The main claim refers to new peptides of formula X-Ala-Arg-Arg-J in which
CC X is H or one or more amino acids and J is OH or one or more amino acids,
CC the peptides containing a total of 3 to 15 amino acids. The present
CC peptide is a specifically claimed example of these new peptides. The
CC peptides are specific inhibitors of protein kinase C isotype zeta, i.e.
CC any subspecies of PKC which contains the specific autoinhibitory
CC pseudosubstrate domain RRGARRWRK (Acc. No. AAR42573). This domain has
CC been found to be perfectly conserved in zeta-PKC variants isolated from a
CC number of different sources, including rat brain. The peptides are
CC usefully therapeutically for treating conditions where the underlying
CC aetiology is associated with zeta-PKC, including tumours,
CC hyperproliferative disorders (e.g. psoriasis) and viral infections (e.g.
CC HIV). (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 30
AAR30753
ID AAR30753 standard; peptide; 3 AA.
XX
AC AAR30753;
XX
DT 26-MAY-1993 (first entry)
XX
DE IgE-mast cell reaction inhibitory peptide prepn. peptide.
XX
KW Physiologically active; basophil; tripeptide.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /note= "N-t-butoxycarbonyl-beta-benzyl-L-Asp"
FT Modified-site 5
FT /note= "N(G)-nitro-L-Arg benzyl ester"
XX
PN JP04360899-A.
XX
PD 14-DEC-1992.
XX
PF 04-JUN-1991; 91JP-00159492.
XX
PR 04-JUN-1991; 91JP-00159492.
XX
PA (TANA) TANABE SEIYAKU CO.
XX
DR WPI; 1993-032711/04.
XX
PT Prepn. of physiologically active penta:peptide - comprises fragment-
PT condensing specified carboxyl protected tri:peptide and amino-protected
PT di:peptide and removing protective gps.
XX
PS Claim; Page 2; 8pp; Japanese.
XX
CC The peptide is used in the prepn. (claimed) of a physiologically active
CC and prevents the IgE - mast cell (basophil) reaction
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 31
AAR48960
ID AAR48960 standard; protein; 3 AA.
XX
AC AAR48960;
XX
DT 25-MAR-2003 (revised)
DT 12-SEP-1994 (first entry)
XX
DE NL4-3 truncated GAG C-terminal peptide.
XX
KW HIV-1; HXB2; antisense; sequence inversion; antisense virus; infection;
KW naturally occuring virus; NOV; translation; replication; infectivity;
KW hepatitis B; HIV-2; SIV; flip-over PCR.
XX
OS Synthetic.
XX
PN WO9403596-A1.
XX
PD 17-FEB-1994.
XX
PF 30-JUL-1993; 93WO-US007179.
XX
PR 30-JUL-1992; 92US-00921104.
XX
PA (UYHA-) UNIV HAWAII.
XX
PI Hu W, Wang J;
XX
DR WPI; 1994-065685/08.
DR N-PSDB; AAQ57688.
XX
PT New antisense viruses and anti:sense-ribozyme viruses - used for treating
PT or preventing viral infections, partic. HIV-1, HIV-2 or SIV infection.
XX
PS Disclosure; Page 108; 167pp; English.
XX
CC This sequence is encoded by a PCR fragment of NL4-3 and represents the C-
CC terminal peptide fragment of the truncated GAG protein. The DNA encoding
CC this fragment was ligated into ClaI/SalI digested pX and the
CC corresponding plasmid was used to produce the antisense virus of the
CC invention. Antisense or truncated RNAs expressed by these viruses bind to
CC the mRNAs expressed by the naturally occuring viruses (NOVs) and prevent
CC the mRNAs from being translated into proteins, thereby preventing the NOV
CC from replicating. The antisense viruses maintain the infectivity of the
CC NOVs, allowing antisense RNAs to reach the mRNAs of the natural viruses.
CC Antisense viruses such as these may be used for treating or preventing a
CC viral infection, particularly HIV-1, HIV-2 or SIV infection or hepatitis
CC B infection. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;

```
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 32
AAR46824
ID   AAR46824 standard; protein; 3 AA.
XX
AC   AAR46824;
XX
DT   25-MAR-2003  (revised)
DT   19-AUG-1994  (first entry)
XX
DE   Phytase derived peptide C-terminal (C phy).
XX
KW   pH 2.5; acid phosphatase; Trichoderma; host; Aspergillus; phytic acid;
KW   phytate degrading enzyme; PDE; removal; inositol hexaphosphoric acid;
KW   plant; feed composition; filtration.
XX
OS   Synthetic.
XX
PN   WO9403612-A1.
XX
PD   17-FEB-1994.
XX
PF   30-JUL-1993;   93WO-FI000310.
XX
PR   31-JUL-1992;   92US-00923724.
XX
PA   (ALKO-) ALKO LTD.
XX
PI   Nevalainen HK,  Paloheimo MT,  Miettinen-Oinonen ASK,  Torkkeli TK;
PI   Cantrell M,  Piddington C,  Rambosek JA,  Turunen MK,  Fagerstroem RB;
XX
DR   WPI; 1994-065700/08.
XX
PT   Compsns. contg. phytate degrading enzymes - obtd. by expression of their
PT   genes in Trichoderma, used partic. for producing animal feed compsns.
XX
PS   Example 4; Page 45; 142pp; English.
XX
CC   The sequences given in AAR46793-824 are peptides derived from the phytase
CC   protein. The phytase protein may be used in the composition of the
CC   invention. The DNA encoding the phytase protein may be introduced into a
CC   Trichoderma host which then expresses it and the protein is collected
CC   from the culture medium. By using Trichoderma as a host for Aspergillus
CC   phytate degrading enzymes such as this, a totally different enzyme
CC   composition compared to that secreted from Aspergillus results. The
CC   enzyme composition can be used for removal of phytic acid or inositol
CC   hexaphosphoric acid from raw material, particularly plant material. The
CC   composition is used in feed compositions for animals. By using
CC   Trichoderma as a source of a composition containing phytate degrading
CC   enzymes some difficult downstream processing problems, eg. filtration,
CC   that occur with similar Aspergillus compositions are avoided and yields
CC   are improved. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 33
AAR44666
ID   AAR44666 standard; peptide; 3 AA.
XX
AC   AAR44666;
XX
DT   20-JAN-1995  (first entry)
XX
DE   Platelet aggregation or adhesion inhibitor - peptide 5.
XX
KW   Platelet aggregation; adhesion; inhibitor; guest; host;
KW   beta-cyclodextrin; protease; resistance; degradation.
XX
OS   Synthetic.
XX
PN   JP06116289-A.
XX
PD   26-APR-1994.
XX
PF   09-OCT-1992;   92JP-00271294.
XX
PR   09-OCT-1992;   92JP-00271294.
XX
PA   (FUJF ) FUJI PHOTO FILM CO LTD.
XX
DR   WPI; 1994-173759/21.
XX
PT   Complex of adhesion peptide in host molecule e.g. beta-cyclodextrin -
PT   useful as platelet aggregation inhibitor which is resistant to protease
PT   degradation in-vivo.
XX
PS   Disclosure; Page 4; 5pp; Japanese.
XX
CC   A peptide complex contg. peptide 5 as guest mol. is useful as a platelet
CC   aggregation or adhesion inhibitor. It is hardly hydrolysed by protease in
CC   vivo and thus maintains its effect for a long period
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 34
AAR63264
ID   AAR63264 standard; peptide; 3 AA.
XX
AC   AAR63264;
XX
DT   25-MAR-2003  (revised)
DT   21-JUL-1995  (first entry)
XX
DE   Thrombin inhibitor peptide, CSAP.
XX
KW   Inhibitor; thrombin; chimeric molecule; fibrin-binding; antibody;
KW   hirudin; monoclonal; 59D8; fibrinopeptide B; coagulation;
KW   coronary stent implantation; adjunctive therapy; fibrinogen; haemorrhage.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Misc-difference 1
FT                   /note= "D form residue"
FT   Modified-site   3
FT                   /note= "opt. modified with -H to give ALD, or CH2Cl to
FT                   give CMK"
XX
PN   WO9425491-A1.
XX
PD   10-NOV-1994.
XX
```

```
PF   03-MAY-1994;   94WO-US004881.
XX
PR   03-MAY-1993;   93US-00058699.
XX
PA   (HARD ) HARVARD COLLEGE.
PA   (UYEM-) UNIV EMORY.
XX
PI   Haber E,  Bode C,  Runge M;
XX
DR   WPI; 1994-358195/44.
XX
PT   Fibrin-binding antibody linked to thrombin inhibitor - useful for
PT   preventing blood coagulation by specifically targetting inhibitor to site
PT   of thrombin activity.
XX
PS   Claim 7; Page 38; 53pp; English.
XX
CC   This sequence represents an inhibitor of thrombin which was used in the
CC   chimeric molecule of the invention. The chimeric molecule further
CC   comprises a fibrin-binding antibody linked to the thrombin inhibitor
CC   through a covalent linkage. The chimeric molecule allows fibrin-specific
CC   antibody targetting of hirudin and other thrombin inhibtors, which is
CC   more potent than thrombin on its own. The fibrin-specific antibody is
CC   pref. the monoclonal antibody, 59D8. The epitope to which 59D8 binds
CC   becomes available only after thrombin cleaves fibrinopeptide B. The
CC   chimeric protein may be used for preventing coagulation of the blood.
CC   Anti-thrombin targetting can be esp. useful in highly thrombogenic
CC   situations such as coronary stent implantation and can be used as an
CC   adjunctive therapy with highly selective thrombolytic agents. The
CC   thrombin inhibitor is localised to sites of thrombin activity by the
CC   antibody which binds to thrombin but does not cross react with uncleaved
CC   fibrinogen. The selectivity of inhibition allows te total amount of
CC   thrombin inhibitor used to be substantially reduced, resulting in a
CC   reduced potential for generalised haemorrhaging. (Updated on 25-MAR-2003
CC   to correct PN field.)
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 35
AAR63263
ID   AAR63263 standard; peptide; 3 AA.
XX
AC   AAR63263;
XX
DT   25-MAR-2003  (revised)
DT   21-JUL-1995  (first entry)
XX
DE   Thrombin inhibitor peptide #4.
XX
KW   Inhibitor; thrombin; chimeric molecule; fibrin-binding; antibody;
KW   hirudin; monoclonal; 59D8; fibrinopeptide B; coagulation;
KW   coronary stent implantation; adjunctive therapy; fibrinogen; haemorrhage.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= H-(D)-Phe
FT   Modified-site   3
FT                   /note= "boroArg-C10H16"
XX
PN   WO9425491-A1.
XX
PD   10-NOV-1994.
XX
PF   03-MAY-1994;   94WO-US004881.
XX
PR   03-MAY-1993;   93US-00058699.
XX
PA   (HARD ) HARVARD COLLEGE.
PA   (UYEM-) UNIV EMORY.
XX
PI   Haber E,  Bode C,  Runge M;
XX
DR   WPI; 1994-358195/44.
XX
PT   Fibrin-binding antibody linked to thrombin inhibitor - useful for
PT   preventing blood coagulation by specifically targetting inhibitor to site
PT   of thrombin activity.
XX
PS   Claim 7; Page 38; 53pp; English.
XX
CC   This sequence represents an inhibitor of thrombin which was used in the
CC   chimeric molecule of the invention. The chimeric molecule further
CC   comprises a fibrin-binding antibody linked to the thrombin inhibitor
CC   through a covalent linkage. The chimeric molecule allows fibrin-specific
CC   antibody targetting of hirudin and other thrombin inhibtors, which is
CC   more potent than thrombin on its own. The fibrin-specific antibody is
CC   pref. the monoclonal antibody, 59D8. The epitope to which 59D8 binds
CC   becomes available only after thrombin cleaves fibrinopeptide B. The
CC   chimeric protein may be used for preventing coagulation of the blood.
CC   Anti-thrombin targetting can be esp. useful in highly thrombogenic
CC   situations such as coronary stent implantation and can be used as an
CC   adjunctive therapy with highly selective thrombolytic agents. The
CC   thrombin inhibitor is localised to sites of thrombin activity by the
CC   antibody which binds to thrombin but does not cross react with uncleaved
CC   fibrinogen. The selectivity of inhibition allows te total amount of
CC   thrombin inhibitor used to be substantially reduced, resulting in a
CC   reduced potential for generalised haemorrhaging. (Updated on 25-MAR-2003
CC   to correct PN field.)
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 36
AAR63261
ID   AAR63261 standard; peptide; 3 AA.
XX
AC   AAR63261;
XX
DT   25-MAR-2003  (revised)
DT   21-JUL-1995  (first entry)
XX
DE   Thrombin inhibitor peptide #2.
XX
KW   Inhibitor; thrombin; chimeric molecule; fibrin-binding; antibody;
KW   hirudin; monoclonal; 59D8; fibrinopeptide B; coagulation;
KW   coronary stent implantation; adjunctive therapy; fibrinogen; haemorrhage.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /label= Boc-(D)-Phe
FT   Modified-site   3
FT                   /note= "boroArg-C10H16"
XX
PN   WO9425491-A1.
XX
```

PD 10-NOV-1994.
XX
PF 03-MAY-1994; 94WO-US004881.
XX
PR 03-MAY-1993; 93US-00058699.
XX
PA (HARD) HARVARD COLLEGE.
PA (UYEM-) UNIV EMORY.
XX
PI Haber E, Bode C, Runge M;
XX
DR WPI; 1994-358195/44.
XX
PT Fibrin-binding antibody linked to thrombin inhibitor - useful for
PT preventing blood coagulation by specifically targetting inhibitor to site
PT of thrombin activity.
XX
PS Claim 7; Page 38; 53pp; English.
XX
CC This sequence represents an inhibitor of thrombin which was used in the
CC chimeric molecule of the invention. The chimeric molecule further
CC comprises a fibrin-binding antibody linked to the thrombin inhibitor
CC through a covalent linkage. The chimeric molecule allows fibrin-specific
CC antibody targetting of hirudin and other thrombin inhibtors, which is
CC more potent than thrombin on its own. The fibrin-specific antibody is
CC pref. the monoclonal antibody, 59D8. The epitope to which 59D8 binds
CC becomes available only after thrombin cleaves fibrinopeptide B. The
CC chimeric protein may be used for preventing coagulation of the blood.
CC Anti-thrombin targetting can be esp. useful in highly thrombogenic
CC situations such as coronary stent implantation and can be used as an
CC adjunctive therapy with highly selective thrombolytic agents. The
CC thrombin inhibitor is localised to sites of thrombin activity by the
CC antibody which binds to thrombin but does not cross react with uncleaved
CC fibrinogen. The selectivity of inhibition allows te total amount of
CC thrombin inhibitor used to be substantially reduced, resulting in a
CC reduced potential for generalised haemorrhaging. (Updated on 25-MAR-2003
CC to correct PN field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 37
AAR63260
ID AAR63260 standard; peptide; 3 AA.
XX
AC AAR63260;
XX
DT 25-MAR-2003 (revised)
DT 21-JUL-1995 (first entry)
XX
DE Thrombin inhibitor peptide #1.
XX
KW Inhibitor; thrombin; chimeric molecule; fibrin-binding; antibody;
KW hirudin; monoclonal; 59D8; fibrinopeptide B; coagulation;
KW coronary stent implantation; adjunctive therapy; fibrinogen;
KW haemorrhaging.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= Ac-(D)-Phe
FT Modified-site 3
FT /note= "boroArg"
XX

PN WO9425491-A1.
XX
PD 10-NOV-1994.
XX
PF 03-MAY-1994; 94WO-US004881.
XX
PR 03-MAY-1993; 93US-00058699.
XX
PA (HARD) HARVARD COLLEGE.
PA (UYEM-) UNIV EMORY.
XX
PI Haber E, Bode C, Runge M;
XX
DR WPI; 1994-358195/44.
XX
PT Fibrin-binding antibody linked to thrombin inhibitor - useful for
PT preventing blood coagulation by specifically targetting inhibitor to site
PT of thrombin activity.
XX
PS Claim 7; Page 38; 53pp; English.
XX
CC This sequence represents an inhibitor of thrombin which was used in the
CC chimeric molecule of the invention. The chimeric molecule further
CC comprises a fibrin-binding antibody linked to the thrombin inhibitor
CC through a covalent linkage. The chimeric molecule allows fibrin-specific
CC antibody targetting of hirudin and other thrombin inhibtors, which is
CC more potent than thrombin on its own. The fibrin-specific antibody is
CC pref. the monoclonal antibody, 59D8. The epitope to which 59D8 binds
CC becomes available only after thrombin cleaves fibrinopeptide B. The
CC chimeric protein may be used for preventing coagulation of the blood.
CC Anti-thrombin targetting can be esp. useful in highly thrombogenic
CC situations such as coronary stent implantation and can be used as an
CC adjunctive therapy with highly selective thrombolytic agents. The
CC thrombin inhibitor is localised to sites of thrombin activity by the
CC antibody which binds to thrombin but does not cross react with uncleaved
CC fibrinogen. The selectivity of inhibition allows te total amount of
CC thrombin inhibitor used to be substantially reduced, resulting in a
CC reduced potential for generalised haemorrhaging. (Updated on 25-MAR-2003
CC to correct PN field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 38
AAR63262
ID AAR63262 standard; peptide; 3 AA.
XX
AC AAR63262;
XX
DT 25-MAR-2003 (revised)
DT 21-JUL-1995 (first entry)
XX
DE Thrombin inhibitor peptide #3.
XX
KW Inhibitor; thrombin; chimeric molecule; fibrin-binding; antibody;
KW hirudin; monoclonal; 59D8; fibrinopeptide B; coagulation;
KW coronary stent implantation; adjunctive therapy; fibrinogen; haemorrhage.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT Modified-site 1
FT /label= H-(D)-Phe
FT Modified-site 3
FT /note= "boroArg-OH"

XX
PN WO9425491-A1.
XX
PD 10-NOV-1994.
XX
PF 03-MAY-1994; 94WO-US004881.
XX
PR 03-MAY-1993; 93US-00058699.
XX
PA (HARD) HARVARD COLLEGE.
PA (UYEM-) UNIV EMORY.
XX
PI Haber E, Bode C, Runge M;
XX
DR WPI; 1994-358195/44.
XX
PT Fibrin-binding antibody linked to thrombin inhibitor - useful for
PT preventing blood coagulation by specifically targetting inhibitor to site
PT of thrombin activity.
XX
PS Claim 7; Page 38; 53pp; English.
XX
CC This sequence represents an inhibitor of thrombin which was used in the
CC chimeric molecule of the invention. The chimeric molecule further
CC comprises a fibrin-binding antibody linked to the thrombin inhibitor
CC through a covalent linkage. The chimeric molecule allows fibrin-specific
CC antibody targetting of hirudin and other thrombin inhibtors, which is
CC more potent than thrombin on its own. The fibrin-specific antibody is
CC pref. the monoclonal antibody, 59D8. The epitope to which 59D8 binds
CC becomes available only after thrombin cleaves fibrinopeptide B. The
CC chimeric protein may be used for preventing coagulation of the blood.
CC Anti-thrombin targetting can be esp. useful in highly thrombogenic
CC situations such as coronary stent implantation and can be used as an
CC adjunctive therapy with highly selective thrombolytic agents. The
CC thrombin inhibitor is localised to sites of thrombin activity by the
CC antibody which binds to thrombin but does not cross react with uncleaved
CC fibrinogen. The selectivity of inhibition allows te total amount of
CC thrombin inhibitor used to be substantially reduced, resulting in a
CC reduced potential for generalised haemorrhaging. (Updated on 25-MAR-2003
CC to correct PN field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 39
AAR51440
ID AAR51440 standard; peptide; 3 AA.
XX
AC AAR51440;
XX
DT 25-MAR-2003 (revised)
DT 27-OCT-1994 (first entry)
XX
DE IGF-1 analogue N-terminal.
XX
KW Insulin-like growth factor; IGF-1; IGF-2; bird; chicken; egg; in ovo;
KW growth; promotion; increase; long R3 IGF-1; LR3 IGF-1.
XX
OS Homo sapiens.
XX
PN WO9406445-A1.
XX
PD 31-MAR-1994.
XX
PF 02-SEP-1993; 93WO-US008279.
XX
PR 17-SEP-1992; 92US-00947035.
XX
PA (EMBR-) EMBREX INC.
PA (GROP-) GROPEP PTY LTD.
PA (USDA) US SEC OF AGRIC.
XX
PI Francis GL, Walton PE, Ballard FJ, Mcmurtry JP, Phelps PV;
XX
DR WPI; 1994-118144/14.
XX
PT Increasing growth of birds - with insulin-like growth factor delivered to
PT the egg before hatching esp. for increasing wt. gain in chickens.
XX
PS Claim 6; Page 37; 45pp; English.
XX
CC Growth of birds is increased by (a) admin. to the bird, in ovo, insulin-
CC like growth factor (IGF)-1 or -2 or their active analogues; (b)
CC incubating to hatch and (c) growing the birds for at least 3 weeks after
CC hatch. IGF-1 analogues used pref. (i) lack 1-5 N-terminal amino aicds or
CC (ii) have Glu(3) absent or replaced by Gly, Gln, Leu, Arg or Lys, partic.
CC Gly and the Thr normally adjacent to Glu can be replaced by Arg or Gly.
CC The IGF-1 analogue pref. comprises the N-terminal given in AAR51439-49.
CC IGF-2 analogues used pref. have Glu(2) or Glu(3) absent or replaced by
CC Gly, Gln, Leu, Arg or Lys, partic. Gly and the Thr normally adjacent to
CC Glu can be replaced by Arg or Gly. The IGF-2 analogue pref. comprises the
CC N-terminal given in AAR51450-53. Esp. Long R3 IGF-1, given in AAR51454 is
CC used. It is the full human IGF-1 sequence with Arg replacing Glu(3) and a
CC 13 amino acid N-terminal extension. (Updated on 25-MAR-2003 to correct PN
CC field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 1 R 1

RESULT 40
AAR61094
ID AAR61094 standard; peptide; 3 AA.
XX
AC AAR61094;
XX
DT 21-MAY-1995 (first entry)
XX
DE ACE-inhibiting tripeptide.
XX
KW ACE; angiotensin converting enzyme inhibitor; fish meat alpha-1000;
KW hypotensive.
XX
OS Synthetic.
XX
PN JP06166697-A.
XX
PD 14-JUN-1994.
XX
PF 01-DEC-1992; 92JP-00343573.
XX
PR 01-DEC-1992; 92JP-00343573.
XX
PA (SENM-) SENMI EKESU KK.
XX
DR WPI; 1994-230661/28.
XX
PT New ACE-inhibiting di-, tri- and tetra:peptide(s) - obtd. by treating
PT fish meat alpha-1000 peptide with ODS resin.
XX
PS Claim 1; Page 2; 10pp; Japanese.

```
XX
CC   A total of ten di-, tri- or tetrapeptides are claimed which have
CC   angiotensin converting enzyme inhibiting activity and which are useful as
CC   hypotensives. The present sequence is one of the ten. The peptides are
CC   obtained by treating fish meat alpha-1000 peptide with ODS resin. This
CC   peptide has an ACE inhibiting IC50 value of 205.6 microMolar
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 41
AAR61092
ID   AAR61092 standard; peptide; 3 AA.
XX
AC   AAR61092;
XX
DT   21-MAY-1995  (first entry)
XX
DE   ACE-inhibiting tripeptide.
XX
KW   ACE; angiotensin converting enzyme inhibitor; fish meat alpha-1000;
KW   hypotensive.
XX
OS   Synthetic.
XX
PN   JP06166697-A.
XX
PD   14-JUN-1994.
XX
PF   01-DEC-1992;   92JP-00343573.
XX
PR   01-DEC-1992;   92JP-00343573.
XX
PA   (SENM-) SENMI EKESU KK.
XX
DR   WPI; 1994-230661/28.
XX
PT   New ACE-inhibiting di-, tri- and tetra:peptide(s) - obtd. by treating
PT   fish meat alpha-1000 peptide with ODS resin.
XX
PS   Claim 1; Page 2; 10pp; Japanese.
XX
CC   A total of ten di-, tri- or tetrapeptides are claimed which have
CC   angiotensin converting enzyme inhibiting activity and which are useful as
CC   hypotensives. The present sequence is one of the ten. The peptides are
CC   obtained by treating fish meat alpha-1000 peptide with ODS resin. This
CC   peptide has an ACE inhibiting IC50 value of 20.0 microMolar
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 42
AAR61093
ID   AAR61093 standard; peptide; 3 AA.
XX
AC   AAR61093;
XX
DT   21-MAY-1995  (first entry)
XX
DE   ACE-inhibiting tripeptide.
XX
KW   ACE; angiotensin converting enzyme inhibitor; fish meat alpha-1000;
KW   hypotensive.
XX
OS   Synthetic.
XX
PN   JP06166697-A.
XX
PD   14-JUN-1994.
XX
PF   01-DEC-1992;   92JP-00343573.
XX
PR   01-DEC-1992;   92JP-00343573.
XX
PA   (SENM-) SENMI EKESU KK.
XX
DR   WPI; 1994-230661/28.
XX
PT   New ACE-inhibiting di-, tri- and tetra:peptide(s) - obtd. by treating
PT   fish meat alpha-1000 peptide with ODS resin.
XX
PS   Claim 1; Page 2; 10pp; Japanese.
XX
CC   A total of ten di-, tri- or tetrapeptides are claimed which have
CC   angiotensin converting enzyme inhibiting activity and which are useful as
CC   hypotensives. The present sequence is one of the ten. The peptides are
CC   obtained by treating fish meat alpha-1000 peptide with ODS resin. This
CC   peptide has an ACE inhibiting IC50 value of 330 microMolar
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 43
AAR73934
ID   AAR73934 standard; peptide; 3 AA.
XX
AC   AAR73934;
XX
DT   14-JUL-1995  (first entry)
XX
DE   Novel tripeptide inhibitor of angiotensin transferase.
XX
KW   angiotensin transferase; inhibitor; hypertension; casein; treatment;
KW   enzymatic hydrolysis.
XX
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "H-Val"
FT   Modified-site   3
FT                   /note= "Tyr-OH"
XX
PN   JP06277091-A.
XX
PD   04-OCT-1994.
XX
PF   26-MAR-1993;   93JP-00092553.
XX
PR   26-MAR-1993;   93JP-00092553.
XX
PA   (NISY ) NIPPON SYNTHETIC CHEM IND CO.
```

```
XX
DR   WPI; 1994-353767/44.
XX
PT   Novel tri-peptide inhibitor of angiotensin converting enzyme - obtained
PT   by enzymatic hydrolysis of protein, e.g. casein.
XX
PS   Claim 1; Page 2; 4pp; Japanese.
XX
CC   AAR73934 is a novel tripeptide useful for the treatment of hypertension.
CC   It inhibits angiotensin transferase and is obtd. by the enzymatic
CC   hydrolysis of proteins using thermolysin, in partic. casein. It is also
CC   useful in the treatment of angina pectoris, ischaemic cardiac
CC   insufficiency, and myocardial infarction
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 44
AAR57458
ID   AAR57458 standard; protein; 3 AA.
XX
AC   AAR57458;
XX
DT   28-FEB-1995  (first entry)
XX
DE   Lactoferrin derived peptide #21.
XX
KW   Lactoferrin; chemical; enzymatic; hydrolysis; antimicrobial; antiseptic;
KW   ischaemic disease.
XX
OS   Mus musculus.
XX
PN   JP06172200-A.
XX
PD   21-JUN-1994.
XX
PF   08-DEC-1992;   92JP-00327738.
XX
PR   08-DEC-1992;   92JP-00327738.
XX
PA   (MORG ) MORINAGA MILK IND CO LTD.
XX
DR   WPI; 1994-238662/29.
XX
PT   Brain protectant for preventing ischaemic diseases without side effects -
PT   comprising 31 specified peptide(s), prepd. by lactoferrin hydrolysis.
XX
PS   Disclosure; Page 9; 11pp; Japanese.
XX
CC   The sequences given in AAR57438-68 represent fragments of lactoferrin
CC   which were derived from the full lenth protein by chemical or enzyme
CC   hydrolysis. These peptides have brain protecting properties, as well as
CC   anti-microbial activity. Compositions containing these peptides may be
CC   prepared with out the addition of antiseptics, and may be administered at
CC   doses of at least 10 mg for parenteral administration and 100 mg for oral
CC   administration. These peptides are stable, heat resistant, water soluble
CC   and may be used for the prevention of ischaemic diseases without side
CC   effects
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 45
AAR57450
ID   AAR57450 standard; protein; 3 AA.
XX
AC   AAR57450;
XX
DT   28-FEB-1995  (first entry)
XX
DE   Lactoferrin derived peptide #13.
XX
KW   Lactoferrin; chemical; enzymatic; hydrolysis; antimicrobial; antiseptic;
KW   ischaemic disease.
XX
OS   Mus musculus.
XX
PN   JP06172200-A.
XX
PD   21-JUN-1994.
XX
PF   08-DEC-1992;   92JP-00327738.
XX
PR   08-DEC-1992;   92JP-00327738.
XX
PA   (MORG ) MORINAGA MILK IND CO LTD.
XX
DR   WPI; 1994-238662/29.
XX
PT   Brain protectant for preventing ischaemic diseases without side effects -
PT   comprising 31 specified peptide(s), prepd. by lactoferrin hydrolysis.
XX
PS   Disclosure; Page 8; 11pp; Japanese.
XX
CC   The sequences given in AAR57438-68 represent fragments of lactoferrin
CC   which were derived from the full lenth protein by chemical or enzyme
CC   hydrolysis. These peptides have brain protecting properties, as well as
CC   anti-microbial activity. Compositions containing these peptides may be
CC   prepared with out the addition of antiseptics, and may be administered at
CC   doses of at least 10 mg for parenteral administration and 100 mg for oral
CC   administration. These peptides are stable, heat resistant, water soluble
CC   and may be used for the prevention of ischaemic diseases without side
CC   effects
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 46
AAR48523
ID   AAR48523 standard; peptide; 3 AA.
XX
AC   AAR48523;
XX
DT   25-MAR-2003  (revised)
DT   10-AUG-1994  (first entry)
XX
DE   Lactoferrin derived peptide #17.
XX
KW   Decomposition; lactoferrin; digestion; enzyme; pepsin; trypsin;
KW   antioxidant; oxidation; inhibitor; vitamin E; ascorbic acid; vitamin A;
KW   beta-carotene; superoxidase dismutase; coenzyme Q; lipid oxidation;
KW   foodstuff; drugs; health food; toiletries; cosmetics.
```

XX
OS Bos taurus.
XX
PN WO9403555-A1.
XX
PD 17-FEB-1994.
XX
PF 04-AUG-1993; 93WO-JP001090.
XX
PR 07-AUG-1992; 92JP-00211335.
XX
PA (MORG) MORINAGA MILK IND CO LTD.
XX
PI Tomita M, Shimamura S, Kawase K, Fukuwatari Y, Takase M;
PI Bellamy WR, Yamauchi K, Wakabayashi H, Tokida Y;
XX
DR WPI; 1994-065650/08.
XX
PT Antioxidant peptide lactoferrin decomposition product - prevents
PT oxidation of lipid(s) in foodstuffs and drugs without affecting their
PT taste.
XX
PS Claim 3; Page 32; 47pp; Japanese.
XX
CC The sequences given in AAR48507-37 are peptides derived by the
CC decomposition of lactoferrin, pref. by digestion with an enzyme, eg.
CC pepsin or trypsin. These peptides may be used in an antioxidant
CC composition which may also contain an oxidation inhibitor such as vitamin
CC E, ascorbic acid, vitamin A, beta-carotene, superoxidase dismutase or
CC coenzyme Q. The antioxidant prevents lipid oxidation in foodstuffs,
CC drugs, health foods, toiletries and cosmetics. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 47
AAR48527
ID AAR48527 standard; peptide; 3 AA.
XX
AC AAR48527;
XX
DT 25-MAR-2003 (revised)
DT 10-AUG-1994 (first entry)
XX
DE Lactoferrin derived peptide #21.
XX
KW Decomposition; lactoferrin; digestion; enzyme; pepsin; trypsin;
KW antioxidant; oxidation; inhibitor; vitamin E; ascorbic acid; vitamin A;
KW beta-carotene; superoxidase dismutase; coenzyme Q; lipid oxidation;
KW foodstuff; drugs; health food; toiletries; cosmetics.
XX
OS Bos taurus.
XX
PN WO9403555-A1.
XX
PD 17-FEB-1994.
XX
PF 04-AUG-1993; 93WO-JP001090.
XX
PR 07-AUG-1992; 92JP-00211335.
XX
PA (MORG) MORINAGA MILK IND CO LTD.
XX
PI Tomita M, Shimamura S, Kawase K, Fukuwatari Y, Takase M;
PI Bellamy WR, Yamauchi K, Wakabayashi H, Tokida Y;
XX
DR WPI; 1994-065650/08.
XX
PT Antioxidant peptide lactoferrin decomposition product - prevents
PT oxidation of lipid(s) in foodstuffs and drugs without affecting their
PT taste.
XX
PS Claim 3; Page 34; 47pp; Japanese.
XX
CC The sequences given in AAR48507-37 are peptides derived by the
CC decomposition of lactoferrin, pref. by digestion with an enzyme, eg.
CC pepsin or trypsin. These peptides may be used in an antioxidant
CC composition which may also contain an oxidation inhibitor such as vitamin
CC E, ascorbic acid, vitamin A, beta-carotene, superoxidase dismutase or
CC coenzyme Q. The antioxidant prevents lipid oxidation in foodstuffs,
CC drugs, health foods, toiletries and cosmetics. (Updated on 25-MAR-2003 to
CC correct PN field.)
XX
SQ Sequence 3 AA;

Query Match 55.6%; Score 5; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.9e+06;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 1 R 1

RESULT 48
AAR48519
ID AAR48519 standard; peptide; 3 AA.
XX
AC AAR48519;
XX
DT 25-MAR-2003 (revised)
DT 10-AUG-1994 (first entry)
XX
DE Lactoferrin derived peptide #13.
XX
KW Decomposition; lactoferrin; digestion; enzyme; pepsin; trypsin;
KW antioxidant; oxidation; inhibitor; vitamin E; ascorbic acid; vitamin A;
KW beta-carotene; superoxidase dismutase; coenzyme Q; lipid oxidation;
KW foodstuff; drugs; health food; toiletries; cosmetics.
XX
OS Bos taurus.
XX
PN WO9403555-A1.
XX
PD 17-FEB-1994.
XX
PF 04-AUG-1993; 93WO-JP001090.
XX
PR 07-AUG-1992; 92JP-00211335.
XX
PA (MORG) MORINAGA MILK IND CO LTD.
XX
PI Tomita M, Shimamura S, Kawase K, Fukuwatari Y, Takase M;
PI Bellamy WR, Yamauchi K, Wakabayashi H, Tokida Y;
XX
DR WPI; 1994-065650/08.
XX
PT Antioxidant peptide lactoferrin decomposition product - prevents
PT oxidation of lipid(s) in foodstuffs and drugs without affecting their
PT taste.
XX
PS Claim 3; Page 30; 47pp; Japanese.
XX
CC The sequences given in AAR48507-37 are peptides derived by the
CC decomposition of lactoferrin, pref. by digestion with an enzyme, eg.
CC pepsin or trypsin. These peptides may be used in an antioxidant
CC composition which may also contain an oxidation inhibitor such as vitamin

```
CC   E, ascorbic acid, vitamin A, beta-carotene, superoxidase dismutase or
CC   coenzyme Q. The antioxidant prevents lipid oxidation in foodstuffs,
CC   drugs, health foods, toiletries and cosmetics. (Updated on 25-MAR-2003 to
CC   correct PN field.)
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 49
AAR58575
ID   AAR58575 standard; peptide; 3 AA.
XX
AC   AAR58575;
XX
DT   26-APR-1995  (first entry)
XX
DE   Angiotensin I converting enzyme inhibitory tripeptide FRP.
XX
KW   angiotensin converting enzyme; inhibitor; food ingredient.
XX
OS   Synthetic.
XX
PN   JP06220088-A.
XX
PD   09-AUG-1994.
XX
PF   22-JAN-1993;   93JP-00025977.
XX
PR   22-JAN-1993;   93JP-00025977.
XX
PA   (ASAH ) ASAHI KASEI KOGYO KK.
XX
DR   WPI; 1994-290911/36.
XX
PT   New tri:peptide(s) - inhibit angiotensin I converting enzyme.
XX
PS   Claim 1; Page 2; 4pp; Japanese.
XX
CC   This is one of thirteen claimed tripeptides (AAR58569-R58581) which
CC   inhibit angiotensin I converting enzyme (with IC50 of 2.9-186.2
CC   micromolar). The tripeptides are incorporated into food, e.g. hamburgers
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 50
AAR58571
ID   AAR58571 standard; peptide; 3 AA.
XX
AC   AAR58571;
XX
DT   26-APR-1995  (first entry)
XX
DE   Angiotensin I converting enzyme inhibitory tripeptide LRY.
XX
KW   angiotensin converting enzyme; inhibitor; food ingredient.
XX
OS   Synthetic.
XX
PN   JP06220088-A.
XX
PD   09-AUG-1994.
XX
PF   22-JAN-1993;   93JP-00025977.
XX
PR   22-JAN-1993;   93JP-00025977.
XX
PA   (ASAH ) ASAHI KASEI KOGYO KK.
XX
DR   WPI; 1994-290911/36.
XX
PT   New tri:peptide(s) - inhibit angiotensin I converting enzyme.
XX
PS   Claim 1; Page 2; 4pp; Japanese.
XX
CC   This is one of thirteen claimed tripeptides (AAR58569-R58581) which
CC   inhibit angiotensin I converting enzyme (with IC50 of 2.9-186.2
CC   micromolar). The tripeptides are incorporated into food, e.g. hamburgers
XX
SQ   Sequence 3 AA;

  Query Match             55.6%;  Score 5;  DB 2;  Length 3;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+06;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


Search completed: January 25, 2006, 18:38:39
Job time : 89.5 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:34:13 ; Search time 12.5 Seconds
(without alignments)
38.487 Million cell updates/sec

Title: US-10-771-242-293
Perfect score: 9
Sequence: 1 RXXXX 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5	55.6	3	3	A22565	R-phycoerythrin al
2	5	55.6	3	3	PQ0010	angiotensin-conver
3	5	55.6	4	1	ECXAA	antho-RFamide neur
4	5	55.6	4	2	A02147	phagocytosis-stimu
5	5	55.6	4	2	D41654	hypothetical prote
6	5	55.6	4	2	I40870	phospholipase C (E
7	5	55.6	4	2	A25844	autho-RF amide neu
8	5	55.6	4	2	I61883	protamine P1 - ora
9	5	55.6	4	2	S47552	ubiquitin - rat
10	5	55.6	4	2	I37013	protamine P1 - Cer
11	5	55.6	4	2	I84439	protamine P1 - sav
12	5	55.6	4	2	A35779	neuropeptide Antho
13	5	55.6	4	2	A60418	FMRFamide - polych
14	5	55.6	4	2	PT0721	T-cell receptor be
15	5	55.6	4	2	ECNK	cardioexcitatory n
16	5	55.6	5	1	HOROHA	proctolin - Americ
17	5	55.6	5	2	JN0862	peptidyl-dipeptida
18	5	55.6	5	2	I40702	primase - Citrobac
19	5	55.6	5	2	A44955	alkanal monooxygen
20	5	55.6	5	2	I39964	ribosomal protein
21	5	55.6	5	2	I39966	ribosomal protein
22	5	55.6	5	2	I39965	ribosomal protein
23	5	55.6	5	2	D60274	major protein anti
24	5	55.6	5	2	B22565	R-phycoerythrin al
25	5	55.6	5	2	F22565	R-phycoerythrin ga
26	5	55.6	5	2	T14910	hypothetical prote
27	5	55.6	5	2	A60803	neuropeptide - sea
28	5	55.6	5	2	S53595	hypothetical prote
29	5	55.6	5	2	PT0295	Ig heavy chain CRD
30	5	55.6	5	2	S62883	seminal plasma pro
31	5	55.6	5	2	PT0513	T-cell receptor be
32	5	55.6	5	2	PT0525	T-cell receptor be
33	5	55.6	5	2	PT0597	T-cell receptor be
34	5	55.6	5	2	PT0672	T-cell receptor be
35	5	55.6	5	2	PT0553	T-cell receptor be
36	5	55.6	5	2	PT0695	T-cell receptor be
37	5	55.6	5	2	PT0577	T-cell receptor be
38	5	55.6	5	2	PT0572	T-cell receptor be
39	5	55.6	5	2	PT0700	T-cell receptor be
40	5	55.6	5	2	A60411	proctolin - Atlant
41	5	55.6	5	2	PT0608	T-cell receptor be
42	5	55.6	5	2	PT0565	T-cell receptor be
43	5	55.6	6	2	A35890	RNA-directed DNA p
44	5	55.6	6	2	A37765	hypothetical prote
45	5	55.6	6	2	C22565	R-phycoerythrin be
46	5	55.6	6	2	PQ0008	angiotensin-conver
47	5	55.6	6	2	A60494	antineoplastic gly
48	5	55.6	6	2	I51434	H4 histone - Afric
49	5	55.6	6	2	I37027	protamine P1 - gor
50	5	55.6	6	2	A11490	pyruvate kinase (E
51	5	55.6	6	2	B56979	collagen alpha 1(I
52	5	55.6	6	2	A19780	transferrin - bovi
53	5	55.6	6	2	B26206	alpha-1,4-glucan-p
54	5	55.6	6	2	I48126	alpha-tubulin - Ch
55	5	55.6	6	2	B33932	Ig mu chain D regi
56	5	55.6	6	2	PT0510	T-cell receptor be
57	5	55.6	6	2	PT0518	T-cell receptor be
58	5	55.6	6	2	PT0618	T-cell receptor be
59	5	55.6	6	2	PT0662	T-cell receptor be
60	5	55.6	6	2	PT0533	T-cell receptor be
61	5	55.6	6	2	PT0568	T-cell receptor be
62	5	55.6	6	2	A41946	T-cell receptor ga
63	5	55.6	6	2	I49424	cytotoxic T-lympho
64	5	55.6	6	2	PC4127	hypothetical 6 pro
65	5	55.6	6	2	A49792	acylaminoacyl-pept
66	5	55.6	6	2	A43129	neuropeptide GNFFR
67	5	55.6	6	4	I79564	hypothetical TCL3
68	5	55.6	7	2	A60224	Met-enkephalin-Arg
69	5	55.6	7	2	PH1408	Ig heavy chain V r
70	5	55.6	7	2	S19630	ribosomal protein
71	5	55.6	7	2	S16364	opacity protein P.
72	5	55.6	7	2	S16365	opacity protein P.
73	5	55.6	7	2	S55548	mcrB protein - Esc
74	5	55.6	7	2	S42407	gramicidin S synth
75	5	55.6	7	2	A28709	phosphonoacetaldeh
76	5	55.6	7	2	S38516	mabinlin II chain
77	5	55.6	7	2	A34818	vicilin 72K chain
78	5	55.6	7	2	PN0150	omega-gliadine 1'
79	5	55.6	7	2	S42620	aggrecan - bovine
80	5	55.6	7	2	E48394	glycoprotein compo
81	5	55.6	7	2	I48105	dihydrofolate redu
82	5	55.6	7	2	B33541	hypothetical prote
83	5	55.6	7	2	E33932	Ig mu chain D regi
84	5	55.6	7	2	PT0602	T-cell receptor be
85	5	55.6	7	2	PT0620	T-cell receptor be
86	5	55.6	7	2	PT0667	T-cell receptor be
87	5	55.6	7	2	PT0655	T-cell receptor be
88	5	55.6	7	2	PT0556	T-cell receptor be
89	5	55.6	7	2	PT0542	T-cell receptor be
90	5	55.6	7	2	PT0567	T-cell receptor be
91	5	55.6	7	2	PT0676	T-cell receptor be
92	5	55.6	7	2	PT0576	T-cell receptor be
93	5	55.6	7	2	PT0581	T-cell receptor be
94	5	55.6	7	2	PT0671	T-cell receptor be
95	5	55.6	7	2	A39690	neural cell adhesi
96	5	55.6	7	2	PH0932	T-cell receptor be
97	5	55.6	7	2	B48394	major fat-globule
98	5	55.6	7	2	A58718	carnocin UI49 - Ca
99	5	55.6	7	2	PQ0777	NADH2 dehydrogenas
100	5	55.6	7	2	S09066	globulin IV alpha
101	5	55.6	7	2	A59489	protein kinase C i
102	5	55.6	7	2	A15398	choline oxidase (E

103	5	55.6	7	2	S45311	microcin C7 - Esch
104	5	55.6	7	2	PC2132	FMRFamide-related
105	5	55.6	7	2	A28340	myomodulin - Calif
106	5	55.6	7	2	S33245	neuromodulatory pe
107	5	55.6	7	2	B44787	calliFMRFamide 11
108	5	55.6	7	2	ECMUCR	catch-relaxing pep
109	5	55.6	7	4	I55382	hypothetical pepti
110	5	55.6	7	4	I56695	hypothetical L2 pr
111	5	55.6	7	4	S15597	orf 4 rara 5'-regi
112	5	55.6	8	1	LFSAME	probable msrA lead
113	5	55.6	8	2	PH1407	Ig heavy chain V r
114	5	55.6	8	2	PC4131	hypothetical prote
115	5	55.6	8	2	T10077	hypothetical prote
116	5	55.6	8	2	S70727	ipgF protein - Shi
117	5	55.6	8	2	B39745	endoglycosylcerami
118	5	55.6	8	2	A39892	P element, P cytot
119	5	55.6	8	2	PT0279	Ig heavy chain CRD
120	5	55.6	8	2	PT0298	Ig heavy chain CRD
121	5	55.6	8	2	PT0323	Ig heavy chain CRD
122	5	55.6	8	2	B45800	serum albumin - do
123	5	55.6	8	2	PH1618	Ig H chain V-D-J r
124	5	55.6	8	2	A35768	T-cell receptor al
125	5	55.6	8	2	PT0522	T-cell receptor be
126	5	55.6	8	2	PT0639	T-cell receptor be
127	5	55.6	8	2	PT0613	T-cell receptor be
128	5	55.6	8	2	PT0653	T-cell receptor be
129	5	55.6	8	2	PT0547	T-cell receptor be
130	5	55.6	8	2	PT0691	T-cell receptor be
131	5	55.6	8	2	B54823	olfactory receptor
132	5	55.6	8	2	I57532	gene TnIslow prote
133	5	55.6	8	2	A37521	R-phycoerythrin ga
134	5	55.6	8	2	S11078	glucose-6-phosphat
135	5	55.6	8	2	PT0030	inulinase (EC 3.2.
136	5	55.6	8	2	PL0162	paramyosin - north
137	5	55.6	8	2	A23967	leucopyrokinin - M
138	5	55.6	8	2	S66646	cardioacceleratory
139	5	55.6	8	2	H41978	calliFMRFamide 8 -
140	5	55.6	8	2	D47393	neuropeptide dalla
141	5	55.6	8	2	S71919	alcohol dehydrogen
142	5	55.6	8	2	A54823	olfactory receptor
143	5	55.6	8	2	I57018	gene Cftr protein
144	5	55.6	8	2	S65381	cytochrome-c oxida
145	5	55.6	8	2	T13818	cytochrome oxidase
146	5	55.6	8	2	S69165	ferredoxin a2 - Ja
147	5	55.6	8	2	A59028	MHC class I histoc
148	5	55.6	8	2	I64832	Ca2+-transporting
149	5	55.6	8	2	A25836	L-serine ammonia-l
150	5	55.6	8	2	T48890	hypothetical prote
151	5	55.6	8	2	B24749	neuropeptide B - b
152	5	55.6	8	4	I55411	hypothetical histo
153	5	55.6	9	2	B45796	dihydrolipoamide S
154	5	55.6	9	2	S59902	glutathione transf
155	5	55.6	9	2	S13889	phosphoenolpyruvat
156	5	55.6	9	2	A28495	conopressin G - co
157	5	55.6	9	2	B28495	conopressin S - co
158	5	55.6	9	2	S35538	ribosomal protein
159	5	55.6	9	2	A44873	caldesmon - rabbit
160	5	55.6	9	2	D24180	fibrinogen beta ch
161	5	55.6	9	2	C24180	fibrinogen beta ch
162	5	55.6	9	2	E28854	fibrinopeptide B -
163	5	55.6	9	2	F28854	fibrinopeptide B -
164	5	55.6	9	2	D28854	fibrinopeptide B -
165	5	55.6	9	2	D58503	translation elonga
166	5	55.6	9	2	JS0302	xenopsin-related p
167	5	55.6	9	2	A60320	xenopsin-related p
168	5	55.6	9	2	S66607	quinoline 2-oxidor
169	5	55.6	9	2	C36730	hutU protein - Kle
170	5	55.6	9	2	T31612	hypothetical prote
171	5	55.6	9	2	S65433	bradykinin - horn
172	5	55.6	9	2	A53797	3',5'-cyclic-GMP p
173	5	55.6	9	2	PT0247	Ig heavy chain CRD
174	5	55.6	9	2	PT0268	Ig heavy chain CRD
175	5	55.6	9	2	PT0299	Ig heavy chain CRD
176	5	55.6	9	2	PT0315	Ig heavy chain CRD
177	5	55.6	9	2	PT0326	Ig heavy chain CRD
178	5	55.6	9	2	C60070	gastrin - domestic
179	5	55.6	9	2	G56978	collagen alpha 1(I
180	5	55.6	9	2	PH1591	Ig H chain V-D-J r
181	5	55.6	9	2	S36850	Ig heavy chain V r
182	5	55.6	9	2	PT0670	T-cell receptor be
183	5	55.6	9	2	PT0634	T-cell receptor be
184	5	55.6	9	2	S78420	ribosomal protein
185	5	55.6	9	2	S39040	lysine-conopressin
186	5	55.6	9	2	S02384	probable membrane
187	5	55.6	9	2	B41983	orf dowstream to b
188	5	55.6	9	2	B39841	dextransucrase (EC
189	5	55.6	9	2	A39841	sucrose 3-glucosyl
190	5	55.6	9	2	A31576	xylose isomerase (
191	5	55.6	9	2	PC2021	oxytocin-related p
192	5	55.6	9	2	B57444	neuropeptide Grb-A
193	5	55.6	9	2	C57444	neuropeptide Grb-A
194	5	55.6	9	2	D57444	neuropeptide Grb-A
195	5	55.6	9	2	A29477	diuretic neuropept
196	5	55.6	9	2	A61620	locustamyotropin I
197	5	55.6	9	2	A41978	calliFMRFamide 1 -
198	5	55.6	9	2	A44787	calliFMRFamide 10
199	5	55.6	9	2	D44787	calliFMRFamide 13
200	5	55.6	9	2	B41978	calliFMRFamide 2 -
201	5	55.6	9	2	C41978	calliFMRFamide 3 -
202	5	55.6	9	2	D41978	calliFMRFamide 4 -
203	5	55.6	9	2	E41978	calliFMRFamide 5 -
204	5	55.6	9	2	F41978	calliFMRFamide 6 -
205	5	55.6	9	2	G41978	calliFMRFamide 7 -
206	5	55.6	9	2	S06375	arginine vasotocin
207	5	55.6	9	2	B61364	vasotocin - common
208	5	55.6	9	2	I50633	c-rel protein - ch
209	5	55.6	9	2	I54379	gene NF2 protein -
210	5	55.6	9	2	A37027	macrophage chemota
211	5	55.6	9	2	S78762	ribosomal protein
212	5	55.6	9	2	PC7076	spectrin alpha cha
213	5	55.6	9	2	I49406	bone gla protein -
214	5	55.6	9	2	S39437	D-amino-acid oxida
215	5	55.6	9	2	A43065	hydroxyproline-3-b
216	5	55.6	9	2	PD0027	pev-tachykinin - p
217	5	55.6	9	2	B60246	ornitho-kinin - ch
218	5	55.6	9	2	A26744	bradykinin-like pe
219	5	55.6	9	2	A61057	Thr-6 bradykinin -
220	5	55.6	9	2	A60579	bradykinin-like pe
221	5	55.6	9	2	B45020	probable minipolyp
222	5	55.6	9	2	A61363	bradykinin - commo
223	5	55.6	9	2	A61358	bradykinin-like pe
224	5	55.6	9	4	S15594	orf 1 rara 5'-regi
225	5	55.6	9	4	S15595	orf 2 rara 5'-regi
226	5	55.6	9	4	I73804	hypothetical E2 pr
227	5	55.6	10	1	XAVI6B	angiotensin-conver
228	5	55.6	10	1	RHPGG	gonadoliberin - pi
229	5	55.6	10	1	RHSHG	gonadoliberin - sh
230	5	55.6	10	1	ECLQ1M	tachykinin I - mig
231	5	55.6	10	1	ECLQ3M	tachykinin III - m
232	5	55.6	10	1	ECLQ4M	tachykinin IV - mi
233	5	55.6	10	1	GMROL2	leucosulfakinin-II
234	5	55.6	10	2	S28055	cytochrome b559 co
235	5	55.6	10	2	S71868	glutathione transf
236	5	55.6	10	2	S39392	calpain (EC 3.4.22
237	5	55.6	10	2	A60624	angiotensin I - Ja
238	5	55.6	10	2	A61131	hydrin 2 - bullfro
239	5	55.6	10	2	A60410	beta-neoendorphin
240	5	55.6	10	2	S39374	mannose receptor -
241	5	55.6	10	2	S66248	processing enzyme,
242	5	55.6	10	2	S51912	hemagglutinin - in
243	5	55.6	10	2	B39517	probable methionin
244	5	55.6	10	2	S59625	beta-galactosidase
245	5	55.6	10	2	S70721	heat shock protein
246	5	55.6	10	2	C39191	hypothetical prote
247	5	55.6	10	2	S42282	parasporal crystal
248	5	55.6	10	2	C38925	seed storage prote

249	5	55.6	10	2	JN0440	peptide-N4-(N-acet
250	5	55.6	10	2	S65432	angiotensin I - ho
251	5	55.6	10	2	PT0230	Ig heavy chain CDR
252	5	55.6	10	2	PT0245	Ig heavy chain CRD
253	5	55.6	10	2	PT0251	Ig heavy chain CRD
254	5	55.6	10	2	PT0284	Ig heavy chain CRD
255	5	55.6	10	2	PT0309	Ig heavy chain CRD
256	5	55.6	10	2	S23370	T-cell receptor al
257	5	55.6	10	2	F49033	T-cell receptor ga
258	5	55.6	10	2	B24736	inhibin beta-B cha
259	5	55.6	10	2	S71948	matrix metalloprot
260	5	55.6	10	2	PH1633	Ig H chain V-D-J r
261	5	55.6	10	2	PH1592	Ig H chain V-D-J r
262	5	55.6	10	2	S36849	Ig heavy chain V r
263	5	55.6	10	2	I48778	small nuclear ribo
264	5	55.6	10	2	PH0807	T-cell receptor al
265	5	55.6	10	2	PT0212	T-cell receptor al
266	5	55.6	10	2	B38887	T-cell receptor ga
267	5	55.6	10	2	D54823	olfactory receptor
268	5	55.6	10	2	A55695	proteoglycan core
269	5	55.6	10	2	PH0933	T-cell receptor be
270	5	55.6	10	2	PH0894	T-cell receptor be
271	5	55.6	10	2	PH0926	T-cell receptor be
272	5	55.6	10	2	PH0923	T-cell receptor be
273	5	55.6	10	2	PH0895	T-cell receptor be
274	5	55.6	10	2	PC4374	telomeric and tetr
275	5	55.6	10	2	I52645	gene B-50 protein
276	5	55.6	10	2	D37397	hypothetical prote
277	5	55.6	10	2	JQ0943	hypothetical 1.3K
278	5	55.6	10	2	S06964	hypothetical prote
279	5	55.6	10	2	S70722	65.4K GTP-binding
280	5	55.6	10	2	A60722	cryptic fimbrial p
281	5	55.6	10	2	PC4442	cytochrome c553 -
282	5	55.6	10	2	S70251	nitrogenase (EC 1.
283	5	55.6	10	2	I40032	trpE protein - Bac
284	5	55.6	10	2	A44646	neurotoxin-associa
285	5	55.6	10	2	I44644	neurotoxin-associa
286	5	55.6	10	2	F41839	ribosomal protein
287	5	55.6	10	2	A40753	aldehyde ferredoxi
288	5	55.6	10	2	PC2044	beta-Kirilowin - M
289	5	55.6	10	2	S38305	lectin GNL2 alpha
290	5	55.6	10	2	D28027	protein P7 - curle
291	5	55.6	10	2	A27617	triose-phosphate i
292	5	55.6	10	2	PS0451	24K protein 4302 -
293	5	55.6	10	2	S19296	16K protein - poul
294	5	55.6	10	2	PX0030	triacylglycerol li
295	5	55.6	10	2	PN0165	triose-phosphate i
296	5	55.6	10	2	A58365	neuropeptide FFRFa
297	5	55.6	10	2	S30348	clotting protein -
298	5	55.6	10	2	B60656	leucosulfakinin II
299	5	55.6	10	2	A43977	FMRFamide-like pro
300	5	55.6	10	2	A42089	transcription fact
301	5	55.6	10	2	C44787	calliFMRFamide 12
302	5	55.6	10	2	A56633	neomyosuppressin -
303	5	55.6	10	2	B56899	serum heterodimer,
304	5	55.6	10	2	A90917	angiotensin precur
305	5	55.6	10	2	A90345	angiotensin precur
306	5	55.6	10	2	S26506	collagen alpha 1(V
307	5	55.6	10	2	C54823	olfactory receptor
308	5	55.6	10	2	A61218	alpha-gliadin 4Ha
309	5	55.6	10	2	B61218	alpha-gliadin 6Ha
310	5	55.6	10	2	T17054	cytochrome-c oxida
311	5	55.6	10	2	T13838	cytochrome-c oxida
312	5	55.6	10	2	T17075	cytochrome-c oxida
313	5	55.6	10	2	T13976	cytochrome-c oxida
314	5	55.6	10	2	T17057	cytochrome-c oxida
315	5	55.6	10	2	T12303	cytochrome-c oxida
316	5	55.6	10	2	T14019	cytochrome-c oxida
317	5	55.6	10	2	T17060	cytochrome-c oxida
318	5	55.6	10	2	T17063	cytochrome-c oxida
319	5	55.6	10	2	T12325	cytochrome-c oxida
320	5	55.6	10	2	T14043	cytochrome-c oxida
321	5	55.6	10	2	T14054	cytochrome-c oxida
322	5	55.6	10	2	T17066	cytochrome-c oxida
323	5	55.6	10	2	T17069	cytochrome-c oxida
324	5	55.6	10	2	T12308	cytochrome-c oxida
325	5	55.6	10	2	T17072	cytochrome-c oxida
326	5	55.6	10	2	T12312	cytochrome-c oxida
327	5	55.6	10	2	T12329	cytochrome-c oxida
328	5	55.6	10	2	T12316	cytochrome-c oxida
329	5	55.6	10	2	T14212	cytochrome-c oxida
330	5	55.6	10	2	T12321	cytochrome-c oxida
331	5	55.6	10	2	T14215	cytochrome-c oxida
332	5	55.6	10	2	T14223	cytochrome-c oxida
333	5	55.6	10	2	T14219	cytochrome-c oxida
334	5	55.6	10	2	C54226	light-harvesting p
335	5	55.6	10	2	PQ0785	NADH2 dehydrogenas
336	5	55.6	10	2	PQ0784	NADH2 dehydrogenas
337	5	55.6	10	2	S39030	lysyl-bradykinin -
338	5	55.6	10	2	S07202	phyllomedulin - tw
339	5	55.6	10	2	G60589	sperm-activating p
340	5	55.6	10	2	B33710	ornithine decarbox
341	5	55.6	10	2	A32543	cardioexcitatory n
342	5	55.6	10	2	S15118	dihydrofolate redu
343	5	55.6	11	1	XASNBA	bradykinin-potenti
344	5	55.6	11	1	XAVIBH	bradykinin-potenti
345	5	55.6	11	1	ECLQ2M	tachykinin II - mi
346	5	55.6	11	1	A60654	substance P - guin
347	5	55.6	11	1	SPHO	substance P - hors
348	5	55.6	11	1	GMROL	leucosulfakinin -
349	5	55.6	11	2	S66196	alcohol dehydrogen
350	5	55.6	11	2	A33917	dihydroorotase (EC
351	5	55.6	11	2	JN0023	substance P - chic
352	5	55.6	11	2	S32575	ribosomal protein
353	5	55.6	11	2	A40693	transgelin - sheep
354	5	55.6	11	2	C53652	rhlR protein - Pse
355	5	55.6	11	2	I54193	Rhesus blood group
356	5	55.6	11	2	D58502	27K bile and gallb
357	5	55.6	11	2	JQ0395	hypothetical prote
358	5	55.6	11	2	S42587	celF protein - Esc
359	5	55.6	11	2	S33782	acetolactate synth
360	5	55.6	11	2	B43669	hypothetical prote
361	5	55.6	11	2	E60691	phycobilisome 8K l
362	5	55.6	11	2	PC2372	58K heat shock pro
363	5	55.6	11	2	B41835	translation elonga
364	5	55.6	11	2	E41476	probable antigen 5
365	5	55.6	11	2	S19301	endo-1,4-beta-xyla
366	5	55.6	11	2	PT0081	protein QA300023 -
367	5	55.6	11	2	T06383	hypothetical prote
368	5	55.6	11	2	S78026	ribosomal protein
369	5	55.6	11	2	PN0167	ribosomal protein
370	5	55.6	11	2	A34135	DNA-binding protei
371	5	55.6	11	2	I52980	glucocerebrosidase
372	5	55.6	11	2	PT0249	Ig heavy chain CRD
373	5	55.6	11	2	PT0250	Ig heavy chain CRD
374	5	55.6	11	2	PT0273	Ig heavy chain CRD
375	5	55.6	11	2	PT0287	Ig heavy chain CRD
376	5	55.6	11	2	PT0302	Ig heavy chain CRD
377	5	55.6	11	2	D56979	collagen alpha 1(I
378	5	55.6	11	2	A33571	follistatin - bovi
379	5	55.6	11	2	PH1632	Ig H chain V-D-J r
380	5	55.6	11	2	PH1600	Ig H chain V-D-J r
381	5	55.6	11	2	PH1583	Ig H chain V-D-J r
382	5	55.6	11	2	PH1584	Ig H chain V-D-J r
383	5	55.6	11	2	PD0442	NIPSNAP2 protein -
384	5	55.6	11	2	PT0217	T-cell receptor be
385	5	55.6	11	2	PT0214	T-cell receptor be
386	5	55.6	11	2	B41946	T-cell receptor ga
387	5	55.6	11	2	C38887	T-cell receptor ga
388	5	55.6	11	2	A49037	TcR gamma V-J regi
389	5	55.6	11	2	PD0441	translation elonga
390	5	55.6	11	2	S53436	beta-D-galactosida
391	5	55.6	11	2	S45386	low density lipopr
392	5	55.6	11	2	S78422	ribosomal protein
393	5	55.6	11	2	PH0929	T-cell receptor be
394	5	55.6	11	2	PH0938	T-cell receptor be

395	5	55.6	11	2	PH0947	T-cell receptor be
396	5	55.6	11	2	PH0903	T-cell receptor be
397	5	55.6	11	2	PH0919	T-cell receptor be
398	5	55.6	11	2	S60294	tubulin 2 beta-3 c
399	5	55.6	11	2	S42449	antl protein - pha
400	5	55.6	11	2	S66606	quinoline 2-oxidor
401	5	55.6	11	2	S04875	nifS protein - Bra
402	5	55.6	11	2	A44755	20alpha-hydroxyste
403	5	55.6	11	2	C59151	protein-tyrosine k
404	5	55.6	11	2	A26120	6-phosphofructokin
405	5	55.6	11	2	S45698	gamma-MSH-like pro
406	5	55.6	11	2	A34662	Achatina cardio-ex
407	5	55.6	11	2	S05002	corazonin - Americ
408	5	55.6	11	2	A60656	perisulfakinin - A
409	5	55.6	11	2	S65395	chemical-sense-rel
410	5	55.6	11	2	I41978	calliFMRFamide 9 -
411	5	55.6	11	2	S33300	probable substance
412	5	55.6	11	2	C37196	bradykinin-potenti
413	5	55.6	11	2	D37196	bradykinin-potenti
414	5	55.6	11	2	E57789	gallbladder stone
415	5	55.6	11	2	S13279	Ile-Ser-bradykinin
416	5	55.6	11	2	A32428	amine oxidase (cop
417	5	55.6	11	2	A61483	pyridoxal kinase (
418	5	55.6	11	2	A59146	conotoxin au5a - c
419	5	55.6	11	2	B59146	conotoxin au5b - c
420	5	55.6	11	2	T12264	cytochrome-c oxida
421	5	55.6	11	2	T12253	cytochrome-c oxida
422	5	55.6	11	2	T12244	cytochrome-c oxida
423	5	55.6	11	2	T12248	cytochrome-c oxida
424	5	55.6	11	2	T17078	cytochrome-c oxida
425	5	55.6	11	2	PU0034	dextransucrase (EC
426	5	55.6	11	2	A48973	glucoamylase A1 (E
427	5	55.6	11	2	S23308	substance P - rain
428	5	55.6	11	2	S23306	substance P - Atla
429	5	55.6	11	2	S07207	Crinia-angiotensin
430	5	55.6	11	2	D61033	ranatachykinin D -
431	5	55.6	11	2	A26930	ermG leader peptid
432	5	55.6	11	2	B26744	megascoliakinin -
433	5	55.6	11	2	A61365	phyllokinin - Rohd
434	5	55.6	11	2	A61033	ranatachykinin A -
435	5	55.6	11	2	A57458	gene Gax protein -
436	5	55.6	11	4	S41909	hypothetical prote
437	5	55.6	11	4	PC2124	aminotransferase c
438	5	55.6	11	4	I52708	ELAV-like neuronal
439	5	55.6	11	4	S19015	hypothetical prote
440	5	55.6	11	4	I54081	retinoic acid rece
441	5	55.6	12	1	A43975	locustamyotropin -
442	5	55.6	12	1	A53709	alpha-conotoxin Im
443	5	55.6	12	1	LFECPE	pyrE leader peptid
444	5	55.6	12	2	A29169	phospholipase A2 (
445	5	55.6	12	2	I64829	gene HEXA protein
446	5	55.6	12	2	C49215	urease (EC 3.9.1.5
447	5	55.6	12	2	C36201	1-aminocyclopropan
448	5	55.6	12	2	PH1675	Ig heavy chain V r
449	5	55.6	12	2	S25056	Ig heavy chain - m
450	5	55.6	12	2	S26557	T-cell receptor be
451	5	55.6	12	2	S26556	T-cell receptor be
452	5	55.6	12	2	S26554	T-cell receptor be
453	5	55.6	12	2	S26555	T-cell receptor be
454	5	55.6	12	2	S26541	T-cell receptor be
455	5	55.6	12	2	A54315	entactin/nidogen -
456	5	55.6	12	2	S36902	Em protein - wheat
457	5	55.6	12	2	S01222	translation elonga
458	5	55.6	12	2	S15815	translation elonga
459	5	55.6	12	2	S65629	protoporphyrinogen
460	5	55.6	12	2	S69123	proton-translocati
461	5	55.6	12	2	S34447	binR protein - Sta
462	5	55.6	12	2	A40763	sucrose-6-phosphat
463	5	55.6	12	2	D28551	hypothetical prote
464	5	55.6	12	2	S49547	hypothetical prote
465	5	55.6	12	2	S36899	ribosomal protein
466	5	55.6	12	2	S67528	napin - rape (frag
467	5	55.6	12	2	S18722	matK protein - bee
468	5	55.6	12	2	A33099	163K exoantigen -
469	5	55.6	12	2	S65730	hemoglobin, extrac
470	5	55.6	12	2	C61308	hemocyanin chain 6
471	5	55.6	12	2	S10624	lipovitellin - Afr
472	5	55.6	12	2	S10626	lipovitellin - Afr
473	5	55.6	12	2	PT0228	Ig heavy chain CDR
474	5	55.6	12	2	PT0255	Ig heavy chain CRD
475	5	55.6	12	2	S57570	T cell receptor V-
476	5	55.6	12	2	S47360	T-cell antigen rec
477	5	55.6	12	2	S47363	T-cell antigen rec
478	5	55.6	12	2	PH1182	T-cell receptor al
479	5	55.6	12	2	PH1174	T-cell receptor al
480	5	55.6	12	2	PH1181	T-cell receptor al
481	5	55.6	12	2	S68402	NAD(+)-glycohydrol
482	5	55.6	12	2	A37933	Ig lambda chain J
483	5	55.6	12	2	C30503	Ig gamma-2b chain
484	5	55.6	12	2	PH1587	Ig H chain V-D-J r
485	5	55.6	12	2	PH1605	Ig H chain V-D-J r
486	5	55.6	12	2	PH1606	Ig H chain V-D-J r
487	5	55.6	12	2	PH1611	Ig H chain V-D-J r
488	5	55.6	12	2	PH1581	Ig H chain V-D-J r
489	5	55.6	12	2	S25039	Ig heavy chain V r
490	5	55.6	12	2	S68271	major urinary prot
491	5	55.6	12	2	PH0802	T-cell receptor al
492	5	55.6	12	2	PH0790	T-cell receptor al
493	5	55.6	12	2	PH0746	T-cell receptor be
494	5	55.6	12	2	A42324	cytochrome P450c27
495	5	55.6	12	2	B39690	neural cell adhesi
496	5	55.6	12	2	PH0930	T-cell receptor be
497	5	55.6	12	2	PH0931	T-cell receptor be
498	5	55.6	12	2	A61332	Na+/K+-exchanging
499	5	55.6	12	2	S21163	NAD ADP-ribosyltra
500	5	55.6	12	2	S29859	gene p10 protein -
501	5	55.6	12	2	S71034	potB protein - Sal
502	5	55.6	12	2	S29479	hypothetical prote
503	5	55.6	12	2	T44420	hypothetical prote
504	5	55.6	12	2	JQ2308	hypothetical 1.4K
505	5	55.6	12	2	JQ2318	hypothetical 1.4K
506	5	55.6	12	2	S17540	agglutinin-III lec
507	5	55.6	12	2	PN0160	ribosomal protein
508	5	55.6	12	2	T46656	hypothetical prote
509	5	55.6	12	2	S16204	6-phosphofructokin
510	5	55.6	12	2	A56878	light yellow cell
511	5	55.6	12	2	E44787	calliMIRFamide 1 -
512	5	55.6	12	2	A33900	hydrin 1 - African
513	5	55.6	12	2	A34858	proteinase E - bla
514	5	55.6	12	2	A49637	MHC class II histo
515	5	55.6	12	2	S43170	kinesin light chai
516	5	55.6	12	2	A60528	insulin-like growt
517	5	55.6	12	2	I46922	gene Bota protein
518	5	55.6	12	2	A33520	inhibitory diffusi
519	5	55.6	12	2	PQ0776	NADH2 dehydrogenas
520	5	55.6	12	2	S07436	tachykinin - Afric
521	5	55.6	12	2	A61360	vespakinin M - hor
522	5	55.6	12	2	A61359	vespakinin X - hor
523	5	55.6	12	2	A44874	proboscipedia - fr
524	5	55.6	12	4	JX0315	aminotransferase c
525	5	55.6	12	4	PC2122	aminotransferase c
526	5	55.6	12	4	PC2123	aminotransferase c
527	5	55.6	12	4	S49073	frame shifted cyta
528	5	55.6	13	1	XAVI9B	angiotensin-conver
529	5	55.6	13	1	UNBO	neurotensin - bovi
530	5	55.6	13	1	UNOPBT	neurotensin - brus
531	5	55.6	13	1	MTCMAD	melanotropin alpha
532	5	55.6	13	1	MTHOAD	melanotropin alpha
533	5	55.6	13	1	JTJG3	tremerogen a-13 -
534	5	55.6	13	2	A38929	glutathione peroxi
535	5	55.6	13	2	A28505	neurotensin [valid
536	5	55.6	13	2	A61067	neurotensin - comm
537	5	55.6	13	2	A53608	neurotensin - guin
538	5	55.6	13	2	A32734	enkephalin precurs
539	5	55.6	13	2	PH1676	Ig heavy chain V r
540	5	55.6	13	2	S78519	ribosomal protein

541	5	55.6	13	2	S48210	collagen alpha 1(V
542	5	55.6	13	2	E39778	lactose phosphotra
543	5	55.6	13	2	PN0122	O1L protein - vacc
544	5	55.6	13	2	A60458	protocatechuate 3,
545	5	55.6	13	2	T08533	hypothetical prote
546	5	55.6	13	2	S22995	hypothetical prote
547	5	55.6	13	2	S12388	argA protein - Sal
548	5	55.6	13	2	PC2371	probable endopepti
549	5	55.6	13	2	S36887	ribosomal protein
550	5	55.6	13	2	S23103	erythronolide synt
551	5	55.6	13	2	S09716	2S albumin large c
552	5	55.6	13	2	H44957	protein P18 - comm
553	5	55.6	13	2	PA0089	protein QF200053 -
554	5	55.6	13	2	E60396	antigen 7H8/2 - ma
555	5	55.6	13	2	PS0443	potassium channel
556	5	55.6	13	2	S52356	hypothetical prote
557	5	55.6	13	2	PT0256	Ig heavy chain CRD
558	5	55.6	13	2	PT0293	Ig heavy chain CRD
559	5	55.6	13	2	PT0304	Ig heavy chain CRD
560	5	55.6	13	2	PT0305	Ig heavy chain CRD
561	5	55.6	13	2	PT0331	Ig heavy chain CRD
562	5	55.6	13	2	PH1316	Ig heavy chain DJ
563	5	55.6	13	2	PH1309	Ig heavy chain DJ
564	5	55.6	13	2	S23640	Ig kappa chain J s
565	5	55.6	13	2	S70441	pancreatic elastas
566	5	55.6	13	2	S47356	T-cell antigen rec
567	5	55.6	13	2	S47357	T-cell antigen rec
568	5	55.6	13	2	S47358	T-cell antigen rec
569	5	55.6	13	2	S47359	T-cell antigen rec
570	5	55.6	13	2	S47365	T-cell antigen rec
571	5	55.6	13	2	S47368	T-cell antigen rec
572	5	55.6	13	2	S47371	T-cell antigen rec
573	5	55.6	13	2	S47372	T-cell antigen rec
574	5	55.6	13	2	S47373	T-cell antigen rec
575	5	55.6	13	2	S47376	T-cell antigen rec
576	5	55.6	13	2	S47377	T-cell antigen rec
577	5	55.6	13	2	S47380	T-cell antigen rec
578	5	55.6	13	2	S47381	T-cell antigen rec
579	5	55.6	13	2	S47382	T-cell antigen rec
580	5	55.6	13	2	S47384	T-cell antigen rec
581	5	55.6	13	2	S47385	T-cell antigen rec
582	5	55.6	13	2	S47388	T-cell antigen rec
583	5	55.6	13	2	S47389	T-cell antigen rec
584	5	55.6	13	2	S47390	T-cell antigen rec
585	5	55.6	13	2	S47392	T-cell antigen rec
586	5	55.6	13	2	S47374	T-cell antigen rec
587	5	55.6	13	2	S23372	T-cell receptor al
588	5	55.6	13	2	S61798	T-cell-specific tr
589	5	55.6	13	2	B47415	mannose-1-phosphat
590	5	55.6	13	2	S10562	zona pellucida-bin
591	5	55.6	13	2	A39836	aggrecan - bovine
592	5	55.6	13	2	B56864	dipeptidyl-peptida
593	5	55.6	13	2	C53275	Ig kappa-1 chain J
594	5	55.6	13	2	E53275	Ig kappa-1 chain J
595	5	55.6	13	2	PH1636	Ig H chain V-D-J r
596	5	55.6	13	2	PH1620	Ig H chain V-D-J r
597	5	55.6	13	2	PH1593	Ig H chain V-D-J r
598	5	55.6	13	2	PH1595	Ig H chain V-D-J r
599	5	55.6	13	2	PH1596	Ig H chain V-D-J r
600	5	55.6	13	2	PH1585	Ig H chain V-D-J r
601	5	55.6	13	2	B26406	Ig kappa chain J r
602	5	55.6	13	2	S22761	Ig lambda-2 chain
603	5	55.6	13	2	PH0788	T-cell receptor al
604	5	55.6	13	2	PH0787	T-cell receptor al
605	5	55.6	13	2	PH0799	T-cell receptor al
606	5	55.6	13	2	PH0805	T-cell receptor al
607	5	55.6	13	2	A47630	Ig kappa chain J r
608	5	55.6	13	2	C47630	Ig kappa chain J r
609	5	55.6	13	2	B47630	Ig kappa chain J r
610	5	55.6	13	2	D47630	Ig kappa chain J r
611	5	55.6	13	2	E47630	Ig kappa chain J r
612	5	55.6	13	2	I51905	collecting duct wa
613	5	55.6	13	2	S54344	glyceraldehyde-3-p
614	5	55.6	13	2	AB0764	his operon leader
615	5	55.6	13	2	I54984	aeg-46.5 protein -
616	5	55.6	13	2	PC1008	40K extracellular
617	5	55.6	13	2	S36668	hypothetical prote
618	5	55.6	13	2	PC2369	unidentified 85K p
619	5	55.6	13	2	S20578	ribosomal protein
620	5	55.6	13	2	S09733	photosystem I prot
621	5	55.6	13	2	JQ2309	hypothetical 1.6K
622	5	55.6	13	2	JQ2319	hypothetical 1.6K
623	5	55.6	13	2	PN0168	phosphopyruvate hy
624	5	55.6	13	2	S32471	lymnaDFamide 1 - g
625	5	55.6	13	2	S32472	lymnaDFamide 2 - g
626	5	55.6	13	2	S32473	lymnaDFamide 3 - g
627	5	55.6	13	2	S32474	lymnaDFamide 4 - g
628	5	55.6	13	2	S32475	lymnaDFamide 5 - g
629	5	55.6	13	2	B61620	locustamyotropin I
630	5	55.6	13	2	I84603	deoxynucleotidyltr
631	5	55.6	13	2	S23638	Ig kappa chain J s
632	5	55.6	13	2	G56046	urinary tract ston
633	5	55.6	13	2	S03879	6-phosphofructokin
634	5	55.6	13	2	S32551	glutathione transf
635	5	55.6	13	2	H33932	Ig kappa chain J r
636	5	55.6	13	2	A33933	Ig kappa chain J r
637	5	55.6	13	2	PC4391	cysteine proteinas
638	5	55.6	13	2	G83988	hypothetical prote
639	5	55.6	13	2	H85575	hypothetical prote
640	5	55.6	13	2	S01043	glutamate-ammonia
641	5	55.6	13	2	I50173	alpha-2 collagen -
642	5	55.6	13	2	JH0460	corticostatic pept
643	5	55.6	13	2	JZVHP1	crabrolin - Europe
644	5	55.6	13	2	A61361	bradykinin-like pe
645	5	55.6	13	4	I70075	glycophorin B (mis
646	5	55.6	14	1	NYPG14	hypothalamic tetra
647	5	55.6	14	1	NTKN1M	alpha-conotoxin MI
648	5	55.6	14	1	LFEBWC	trp operon leader
649	5	55.6	14	1	LFEBWT	trp operon leader
650	5	55.6	14	1	LFECW	trp operon leader
651	5	55.6	14	1	LFECFS	pheST operon leade
652	5	55.6	14	1	BSTD	bombesin - fire-be
653	5	55.6	14	2	A47421	leukotriene B-4 12
654	5	55.6	14	2	A33798	D-amino-acid oxida
655	5	55.6	14	2	C40944	hypothetical prote
656	5	55.6	14	2	A01250	angiotensin precur
657	5	55.6	14	2	B61309	lutropin beta chai
658	5	55.6	14	2	A58963	alpha-conotoxin Cn
659	5	55.6	14	2	PH1677	Ig heavy chain V r
660	5	55.6	14	2	PH1705	Ig heavy chain V r
661	5	55.6	14	2	I51432	histone H4-1 precu
662	5	55.6	14	2	A49018	myosin heavy chain
663	5	55.6	14	2	T46634	acyl carrier prote
664	5	55.6	14	2	S50900	chlorophyll a/b-bi
665	5	55.6	14	2	E90858	trp operon leader
666	5	55.6	14	2	B85761	trp operon leader
667	5	55.6	14	2	F90931	pheST operon leade
668	5	55.6	14	2	B85780	pheST operon leade
669	5	55.6	14	2	S60353	amylopullulanase -
670	5	55.6	14	2	A58503	kidney and bladder
671	5	55.6	14	2	PL0142	carbon-monoxide de
672	5	55.6	14	2	PS0371	hypothetical prote
673	5	55.6	14	2	S36892	ribosomal protein
674	5	55.6	14	2	S29789	hypothetical prote
675	5	55.6	14	2	PA0111	protein QA100054 -
676	5	55.6	14	2	S09721	2S albumin small c
677	5	55.6	14	2	PN0151	omega-gliadine 2'
678	5	55.6	14	2	PA0104	protein QF200070 -
679	5	55.6	14	2	B34135	DNA-binding protei
680	5	55.6	14	2	PC1215	homeotic protein E
681	5	55.6	14	2	S23376	collagen alpha cha
682	5	55.6	14	2	PL0152	metal-binding prot
683	5	55.6	14	2	E61308	hemocyanin chain 3
684	5	55.6	14	2	S38307	DEB-A protein - fr
685	5	55.6	14	2	S11074	alcohol dehydrogen
686	5	55.6	14	2	I54284	C1-inhibitor - hum

687	5	55.6	14	2	I64815	carbonic anhydrase
688	5	55.6	14	2	PT0223	Ig heavy chain CDR
689	5	55.6	14	2	PT0232	Ig heavy chain CRD
690	5	55.6	14	2	PT0252	Ig heavy chain CRD
691	5	55.6	14	2	PT0254	Ig heavy chain CRD
692	5	55.6	14	2	PT0294	Ig heavy chain CRD
693	5	55.6	14	2	PH1347	Ig heavy chain DJ
694	5	55.6	14	2	PH1327	Ig heavy chain DJ
695	5	55.6	14	2	PH1332	Ig heavy chain DJ
696	5	55.6	14	2	PH1311	Ig heavy chain DJ
697	5	55.6	14	2	PH1321	Ig heavy chain DJ
698	5	55.6	14	2	PH1305	Ig heavy chain DJ
699	5	55.6	14	2	PH1306	Ig heavy chain DJ
700	5	55.6	14	2	S23639	Ig kappa chain J s
701	5	55.6	14	2	PH1763	T cell receptor al
702	5	55.6	14	2	PH1767	T cell receptor al
703	5	55.6	14	2	S57569	T cell receptor V-
704	5	55.6	14	2	S23369	T-cell receptor al
705	5	55.6	14	2	PH0135	T-cell receptor be
706	5	55.6	14	2	B28018	very late antigen-
707	5	55.6	14	2	S58426	spermadhesin AWN h
708	5	55.6	14	2	F48394	glycoprotein compo
709	5	55.6	14	2	A43847	Ig mu chain V regi
710	5	55.6	14	2	C44823	synaptosomal-assoc
711	5	55.6	14	2	A61032	troponin T, cardia
712	5	55.6	14	2	JS0272	hypothetical 1.5K
713	5	55.6	14	2	PH1625	Ig H chain V-D-J r
714	5	55.6	14	2	PH1626	Ig H chain V-D-J r
715	5	55.6	14	2	PH1627	Ig H chain V-D-J r
716	5	55.6	14	2	PH1628	Ig H chain V-D-J r
717	5	55.6	14	2	PH1614	Ig H chain V-D-J r
718	5	55.6	14	2	PH1639	Ig H chain V-D-J r
719	5	55.6	14	2	PH1617	Ig H chain V-D-J r
720	5	55.6	14	2	PH1623	Ig H chain V-D-J r
721	5	55.6	14	2	PH1586	Ig H chain V-D-J r
722	5	55.6	14	2	PH1594	Ig H chain V-D-J r
723	5	55.6	14	2	PH1597	Ig H chain V-D-J r
724	5	55.6	14	2	PH1608	Ig H chain V-D-J r
725	5	55.6	14	2	PH0792	T-cell receptor al
726	5	55.6	14	2	PH0804	T-cell receptor al
727	5	55.6	14	2	PT0210	T-cell receptor al
728	5	55.6	14	2	PH0765	T-cell receptor be
729	5	55.6	14	2	PH0762	T-cell receptor be
730	5	55.6	14	2	PH0755	T-cell receptor be
731	5	55.6	14	2	D35141	T-cell receptor de
732	5	55.6	14	2	C35141	T-cell receptor de
733	5	55.6	14	2	E35141	T-cell receptor de
734	5	55.6	14	2	S65392	cytochrome-c oxida
735	5	55.6	14	2	PH0945	T-cell receptor be
736	5	55.6	14	2	PH0915	T-cell receptor be
737	5	55.6	14	2	C48394	major fat-globule
738	5	55.6	14	2	G33160	H+-transporting tw
739	5	55.6	14	2	A54370	inorganic diphosph
740	5	55.6	14	2	AG0705	phenylalanyl-tRNA
741	5	55.6	14	2	PH1471	T-cell receptor be
742	5	55.6	14	2	I56388	Km(r) protein - Es
743	5	55.6	14	2	A41589	25K elastin-bindin
744	5	55.6	14	2	S58862	botulinum neurotox
745	5	55.6	14	2	S58866	botulinum neurotox
746	5	55.6	14	2	PA0045	porin porl - Arabi
747	5	55.6	14	2	PT0026	calotropin DI - mu
748	5	55.6	14	2	PN0147	omega-gliadine 1 a
749	5	55.6	14	2	A61306	ribonuclease M (EC
750	5	55.6	14	2	S45655	cathepsin L (EC 3.
751	5	55.6	14	2	A61308	hemocyanin chain 2
752	5	55.6	14	2	G61308	hemocyanin chain 3
753	5	55.6	14	2	D61308	hemocyanin chain 5
754	5	55.6	14	2	A48389	leiurutoxin III -
755	5	55.6	14	2	A56632	neosulfakinin-II -
756	5	55.6	14	2	S12904	protein kinase (EC
757	5	55.6	14	2	PL0040	glycogen phosphory
758	5	55.6	14	2	JH0516	insulin-like growt
759	5	55.6	14	2	D45474	thrombospondin 2 -
760	5	55.6	14	2	S36678	dodecenoyl-CoA Del
761	5	55.6	14	2	S68095	calcium-binding pr
762	5	55.6	14	2	H83778	hypothetical prote
763	5	55.6	14	2	AF0296	phenylalanyl-tRNA
764	5	55.6	14	2	S29878	Na+/K+-exchanging
765	5	55.6	14	2	S27140	hypothetical prote
766	5	55.6	14	2	JH0328	probursin tetradec
767	5	55.6	14	2	A42473	ermK leader peptid
768	5	55.6	14	2	A44515	Trp EG leader pept
769	5	55.6	14	2	A61362	bradykinin-like pe
770	5	55.6	14	2	A32654	fibrinopeptide A -
771	5	55.6	14	4	I52618	hemoglobin beta ch
772	5	55.6	14	4	S00843	hypothetical prote
773	5	55.6	15	1	NTKNAG	alpha-conotoxin GI
774	5	55.6	15	1	LFTWL	leu leader peptide
775	5	55.6	15	2	PA0041	plastoquinol-plast
776	5	55.6	15	2	S42741	ubiquinol-cytochro
777	5	55.6	15	2	B61243	dimethylaniline mo
778	5	55.6	15	2	S14749	3-dehydroquinase -
779	5	55.6	15	2	S21241	oligo-1,6-glucosid
780	5	55.6	15	2	S24159	leukocyte elastase
781	5	55.6	15	2	S57201	basic proteinase I
782	5	55.6	15	2	PC2215	fibrinogenolytic p
783	5	55.6	15	2	A47146	topoisomerase I -
784	5	55.6	15	2	A60834	angiotensin I prec
785	5	55.6	15	2	I49420	placental lactogen
786	5	55.6	15	2	S26525	T-cell receptor al
787	5	55.6	15	2	S26518	T-cell receptor al
788	5	55.6	15	2	S26517	T-cell receptor al
789	5	55.6	15	2	A38304	heterogeneous ribo
790	5	55.6	15	2	S36888	ribosomal protein
791	5	55.6	15	2	I58116	Dp116 - human
792	5	55.6	15	2	I46512	troponin - rabbit
793	5	55.6	15	2	I49407	placental calcium-
794	5	55.6	15	2	I29501	fibrinopeptide A -
795	5	55.6	15	2	P29501	fibrinopeptide A -
796	5	55.6	15	2	JP0101	fibrinogen alpha c
797	5	55.6	15	2	PQ0017	terminal protein -
798	5	55.6	15	2	S59492	formate dehydrogen
799	5	55.6	15	2	JN0730	hypothetical 1.7K
800	5	55.6	15	2	C48401	ribosomal protein
801	5	55.6	15	2	C41383	32K variable histo
802	5	55.6	15	2	S29386	nigerythrin - Desu
803	5	55.6	15	2	A56786	pimeloyl-CoA synth
804	5	55.6	15	2	PA0018	photosystem I 9K p
805	5	55.6	15	2	PA0029	protein QA100012 -
806	5	55.6	15	2	PA0020	protein QA100028 -
807	5	55.6	15	2	PA0024	protein QA300050 -
808	5	55.6	15	2	PQ0193	stylar glycoprotei
809	5	55.6	15	2	PN0148	omega-gliadine 3 -
810	5	55.6	15	2	PS0251	15K protein 5106 -
811	5	55.6	15	2	PS0208	20K protein 5403 -
812	5	55.6	15	2	PC4268	alpha-globulin - r
813	5	55.6	15	2	PC4269	unidentified QR310
814	5	55.6	15	2	PA0057	adenylate isopente
815	5	55.6	15	2	PA0076	fructose-bisphosph
816	5	55.6	15	2	PA0051	protein QF200016 -
817	5	55.6	15	2	PA0060	protein QF200037 -
818	5	55.6	15	2	S71300	ICL3 protein - Par
819	5	55.6	15	2	G60977	protein 425 - Cali
820	5	55.6	15	2	C36198	T-cell receptor be
821	5	55.6	15	2	B32800	hypothetical prote
822	5	55.6	15	2	PT0222	Ig heavy chain CDR
823	5	55.6	15	2	PH1314	Ig heavy chain DJ
824	5	55.6	15	2	PH1342	Ig heavy chain DJ
825	5	55.6	15	2	PH1329	Ig heavy chain DJ
826	5	55.6	15	2	PH1310	Ig heavy chain DJ
827	5	55.6	15	2	PH1788	T cell receptor al
828	5	55.6	15	2	S57584	T cell receptor V-
829	5	55.6	15	2	S57577	T cell receptor V-
830	5	55.6	15	2	G41299	T-cell receptor al
831	5	55.6	15	2	PH0136	T-cell receptor be
832	5	55.6	15	2	G49255	T-cell receptor be

833	5	55.6	15	2	C24687	T-cell receptor be
834	5	55.6	15	2	D28587	T-cell receptor be
835	5	55.6	15	2	F28587	T-cell receptor be
836	5	55.6	15	2	S51735	T-cell receptor be
837	5	55.6	15	2	B49655	T-cell-receptor be
838	5	55.6	15	2	A61247	urogenital tumor m
839	5	55.6	15	2	S66215	cartilage oligomer
840	5	55.6	15	2	E56978	collagen alpha 2(X
841	5	55.6	15	2	B33527	fructose-2,6-bisph
842	5	55.6	15	2	PQ0074	T-cell receptor be
843	5	55.6	15	2	S30608	translation elonga
844	5	55.6	15	2	F44823	synaptosomal-assoc
845	5	55.6	15	2	I78838	flt3 ligand isofor
846	5	55.6	15	2	A27504	histone H2A - mous
847	5	55.6	15	2	PH1613	Ig H chain V-D-J r
848	5	55.6	15	2	PH1616	Ig H chain V-D-J r
849	5	55.6	15	2	PH1590	Ig H chain V-D-J r
850	5	55.6	15	2	PH1610	Ig H chain V-D-J r
851	5	55.6	15	2	PH1582	Ig H chain V-D-J r
852	5	55.6	15	2	PH1378	T antigen variant
853	5	55.6	15	2	PH1377	T antigen variant
854	5	55.6	15	2	PH0775	T-cell receptor al
855	5	55.6	15	2	PH0779	T-cell receptor al
856	5	55.6	15	2	PH0780	T-cell receptor al
857	5	55.6	15	2	PH0806	T-cell receptor al
858	5	55.6	15	2	PH0797	T-cell receptor al
859	5	55.6	15	2	PH0760	T-cell receptor be
860	5	55.6	15	2	PH0772	T-cell receptor be
861	5	55.6	15	2	PH0770	T-cell receptor be
862	5	55.6	15	2	PH0764	T-cell receptor be
863	5	55.6	15	2	B35141	T-cell receptor de
864	5	55.6	15	2	G35141	T-cell receptor de
865	5	55.6	15	2	A35141	T-cell receptor de
866	5	55.6	15	2	E49037	TcR delta chain V-
867	5	55.6	15	2	S71306	heat shock protein
868	5	55.6	15	2	C31409	protein 425 - rat
869	5	55.6	15	2	S21293	RIP protein - rat
870	5	55.6	15	2	A45096	thyrotropin-releas
871	5	55.6	15	2	PL0109	complement factor
872	5	55.6	15	2	S72432	epoxypropan isomer
873	5	55.6	15	2	PT0091	H+-transporting tw
874	5	55.6	15	2	PT0095	H+-transporting tw
875	5	55.6	15	2	S71396	dihydropyrimidine
876	5	55.6	15	2	PT0092	NADH2 dehydrogenas
877	5	55.6	15	2	AF0832	phe leader peptide
878	5	55.6	15	2	C56979	collagen alpha 1(I
879	5	55.6	15	2	H56978	collagen alpha 1(X
880	5	55.6	15	2	PH1455	T-cell receptor al
881	5	55.6	15	2	PH1443	T-cell receptor al
882	5	55.6	15	2	PH1441	T-cell receptor al
883	5	55.6	15	2	S02381	probable membrane
884	5	55.6	15	2	A60929	dichloromethane de
885	5	55.6	15	2	PN0629	integration host f
886	5	55.6	15	2	S67918	serine proteinase
887	5	55.6	15	2	PL0143	carbon-monoxide de
888	5	55.6	15	2	B60929	dichloromethane de
889	5	55.6	15	2	S21411	nodulation protein
890	5	55.6	15	2	A41338	isocitrate lyase (
891	5	55.6	15	2	S71920	proteinase ECP 32
892	5	55.6	15	2	A35389	urease (EC 3.5.1.5
893	5	55.6	15	2	T03000	GTP-binding protei
894	5	55.6	15	2	S33781	acetolactate synth
895	5	55.6	15	2	C43334	orf3 3' to aaDR -
896	5	55.6	15	2	B41868	hypothetical prote
897	5	55.6	15	2	A17340	ribonucleoside-dip
898	5	55.6	15	2	S36891	ribosomal protein
899	5	55.6	15	2	S36896	ribosomal protein
900	5	55.6	15	2	S36889	ribosomal protein
901	5	55.6	15	2	S59489	steroid monooxygen
902	5	55.6	15	2	A40634	orf19 3' of eryK -
903	5	55.6	15	2	T09463	ribosomal protein
904	5	55.6	15	2	S62641	porphobilinogen sy
905	5	55.6	15	2	A56970	GLYMA1 - soybean (
906	5	55.6	15	2	PS0450	23K protein 4307 -
907	5	55.6	15	2	PS0212	29K protein 4228 -
908	5	55.6	15	2	PA0064	cytochrome C 1 - f
909	5	55.6	15	2	PA0087	cytochrome c2 - fu
910	5	55.6	15	2	PA0075	fructose-bisphosph
911	5	55.6	15	2	PA0102	fructose-bisphosph
912	5	55.6	15	2	PA0062	fumarate hydratase
913	5	55.6	15	2	A49177	22K protein p1, mi
914	5	55.6	15	2	A08416	lombricine kinase
915	5	55.6	15	2	A61612	allatostatin - tob
916	5	55.6	15	2	A36527	juvenile-hormone e
917	5	55.6	15	2	PT0090	alpha-glucosidase
918	5	55.6	15	2	I50503	agrin - electric r
919	5	55.6	15	2	A26212	carboxypeptidase B
920	5	55.6	15	2	A32921	beaded-chain filam
921	5	55.6	15	2	A49480	major immunophilin
922	5	55.6	15	2	A32971	heparin-binding le
923	5	55.6	15	2	B45115	peptidylprolyl iso
924	5	55.6	15	2	B45133	casein kinase II (
925	5	55.6	15	2	A35417	28K serine protein
926	5	55.6	15	2	B45474	thrombospondin 2 -
927	5	55.6	15	2	A53085	lipid transfer pro
928	5	55.6	15	2	I46909	voltage-dependent
929	5	55.6	15	2	S62675	collagen type I -
930	5	55.6	15	2	PT0096	pyruvate dehydroge
931	5	55.6	15	2	PT0094	succinate dehydrog
932	5	55.6	15	2	I67525	CD33 antigen homol
933	5	55.6	15	2	A31902	bone acidic glycop
934	5	55.6	15	2	A56963	acid phosphatase (
935	5	55.6	15	2	S29175	D-galactose-bindin
936	5	55.6	15	2	PQ0780	NADH2 dehydrogenas
937	5	55.6	15	2	S04586	NADH2 dehydrogenas
938	5	55.6	15	2	PU0025	ubiquinol-cytochro
939	5	55.6	15	2	A28497	neurotensin-relate
940	5	55.6	15	2	D48648	hypothetical leade
941	5	55.6	15	2	I40665	ilvBN leader pepti
942	5	55.6	15	2	JT0610	leukocyte chemoatt
943	5	55.6	15	4	I38335	hypothetical TEL/M
944	5	55.6	15	4	I52698	hypothetical THRA1
945	5	55.6	15	4	I38336	hypothetical TEL/M
946	5	55.6	16	1	A49761	locustapyrokinin -
947	5	55.6	16	1	MTDFBS	melanotropin beta
948	5	55.6	16	1	LFECH	his operon leader
949	5	55.6	16	2	B42324	cytochrome P450c27
950	5	55.6	16	2	S03405	hydrogenase (EC 1.
951	5	55.6	16	2	C45133	casein kinase II (
952	5	55.6	16	2	S65520	phospholipase A2 (
953	5	55.6	16	2	S10807	protein kinase C i
954	5	55.6	16	2	S10809	protein kinase C i
955	5	55.6	16	2	F44908	chitinase (EC 3.2.
956	5	55.6	16	2	A60551	leukocyte elastase
957	5	55.6	16	2	A44413	proteasome endopep
958	5	55.6	16	2	A59042	alpha-conotoxin Ep
959	5	55.6	16	2	JH0517	insulin-like growt
960	5	55.6	16	2	D49021	Ig heavy chain J7
961	5	55.6	16	2	A28144	ribosomal protein
962	5	55.6	16	2	A27803	myosin light chain
963	5	55.6	16	2	A29501	fibrinopeptide A -
964	5	55.6	16	2	B24180	fibrinogen alpha c
965	5	55.6	16	2	A24180	fibrinogen alpha c
966	5	55.6	16	2	B28854	fibrinopeptide A -
967	5	55.6	16	2	C28854	fibrinopeptide A -
968	5	55.6	16	2	A28854	fibrinopeptide A -
969	5	55.6	16	2	G29501	fibrinopeptide A -
970	5	55.6	16	2	H29501	fibrinopeptide A -
971	5	55.6	16	2	CQRT	cerebellin - rat
972	5	55.6	16	2	B25979	nodulation protein
973	5	55.6	16	2	A45133	casein kinase II (
974	5	55.6	16	2	I78533	gene agoute protei
975	5	55.6	16	2	G45681	orf 61.1 - phage T
976	5	55.6	16	2	A39109	hypothetical prote
977	5	55.6	16	2	S51057	ribosomal protein
978	5	55.6	16	2	S22677	heat shock protein

979	5	55.6	16	2	A53337	regulatory protein
980	5	55.6	16	2	T37075	hypothetical prote
981	5	55.6	16	2	S51610	hypothetical prote
982	5	55.6	16	2	A34053	deoxyribonuclease
983	5	55.6	16	2	B35491	retinol-binding pr
984	5	55.6	16	2	PT0224	Ig heavy chain CDR
985	5	55.6	16	2	PT0282	Ig heavy chain CDR
986	5	55.6	16	2	PT0296	Ig heavy chain CDR
987	5	55.6	16	2	PH1346	Ig heavy chain DJ
988	5	55.6	16	2	PH1351	Ig heavy chain DJ
989	5	55.6	16	2	PH1317	Ig heavy chain DJ
990	5	55.6	16	2	PH1302	Ig heavy chain DJ
991	5	55.6	16	2	PH1778	T cell receptor al
992	5	55.6	16	2	PH1781	T cell receptor al
993	5	55.6	16	2	PH1782	T cell receptor al
994	5	55.6	16	2	PH1790	T cell receptor al
995	5	55.6	16	2	PH1771	T cell receptor al
996	5	55.6	16	2	S57517	T cell receptor be
997	5	55.6	16	2	F41299	T-cell receptor al
998	5	55.6	16	2	PH0137	T-cell receptor be
999	5	55.6	16	2	G49039	T-cell receptor be
1000	5	55.6	16	2	A28587	T-cell receptor be

ALIGNMENTS

RESULT 1
A22565
R-phycoerythrin alpha-1 chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: A22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601; PMID:3886644
A;Accession: A22565
A;Molecule type: protein
A;Residues: 1-3 <KLO>
A;Cross-references: UNIPARC:UPI000017CE9A

Query Match 55.6%; Score 5; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 2
PQ0010
angiotensin-converting enzyme inhibitor (FLP-3) - common fig
N;Alternate names: ficus latex peptide 3
C;Species: Ficus carica (common fig)
C;Date: 15-Jun-2001 #sequence_revision 15-Jun-2001 #text_change 15-Jun-2001
C;Accession: PQ0010
R;Maruyama, S.; Miyoshi, S.; Tanaka, H.
Agric. Biol. Chem. 53, 2763-2767, 1989
A;Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.
A;Reference number: PQ0008
A;Accession: PQ0010
A;Molecule type: protein
A;Residues: 1-3 <MAR>
A;Cross-references: UNIPARC:UPI000011E971
A;Experimental source: latex
C;Keywords: angiotensin-converting enzyme inhibitor

Query Match 55.6%; Score 5; DB 3; Length 3;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 3
ECXAA
antho-RFamide neuropeptide - sea anemone (Anthopleura elegantissima)
C;Species: Anthopleura elegantissima
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C;Accession: A26666
R;Grimmelikhuijzen, C.J.P.; Graff, D.
Proc. Natl. Acad. Sci. U.S.A. 83, 9817-9821, 1986
A;Title: Isolation of <Glu-Gly-Arg-Phe-NH2 (Antho-RFamide), a neuropeptide from sea anem
A;Reference number: A26666; MUID:87092339; PMID:2879288
A;Accession: A26666
A;Molecule type: protein
A;Residues: 1-4 <GRI>
A;Cross-references: UNIPROT:P10419; UNIPARC:UPI00001733AF
C;Comment: The function of this peptide is not known but it could act as a transmitter a
C;Comment: Synthetic and natural peptides had identical properties.
C;Superfamily: RFamide neuropeptide
C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;4/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 55.6%; Score 5; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 4
A02147
phagocytosis-stimulating peptide (tuftsin) - human
C;Species: Homo sapiens (man)
C;Date: 31-Mar-1991 #sequence_revision 31-Mar-1991 #text_change 09-Jul-2004
C;Accession: A02147
R;Nishioka, K.; Constantopoulos, A.; Satoh, P.S.; Najjar, V.A.
Biochem. Biophys. Res. Commun. 47, 172-179, 1972
A;Title: The characteristics, isolation and synthesis of the phagocytosis stimulating pe
A;Reference number: A02147; MUID:72187087; PMID:4112769
A;Accession: A02147
A;Molecule type: protein
A;Residues: 1-4 <NIS>
A;Cross-references: UNIPROT:P01858; UNIPARC:UPI00001377C2
A;Note: a peptide having the same structure, physical properties, and biological activit
R;Fidalgo, B.V.; Najjar, V.A.
Biochemistry 6, 3386-3392, 1967
A;Reference number: A37502; MUID:68091045; PMID:4169272
A;Contents: annotation; immunoglobulin class
C;Comment: An IgG (called leucokinin) binds reversibly to the cell membrane of neutrophi
n is essential for maximum stimulation of the phagocytic activity of neutrophils.
C;Superfamily: immunoglobulin C region; immunoglobulin homology

Query Match 55.6%; Score 5; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

RESULT 5
D41654
hypothetical protein (sodC 5' region) - Haemophilus parainfluenzae (fragment)
C;Species: Haemophilus parainfluenzae
C;Date: 12-Jun-1992 #sequence_revision 12-Jun-1992 #text_change 24-Feb-1995
C;Accession: D41654

R;Kroll, J.S.; Langford, P.R.; Loynds, B.M.
J. Bacteriol. 173, 7449-7457, 1991
A;Title: Copper-zinc superoxide dismutase of Haemophilus influenzae and Haemophilus para
A;Reference number: A41654; MUID:92041655; PMID:1938942
A;Accession: D41654
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-4 <KRO>
A;Cross-references: UNIPARC:UPI000017AB26

Query Match 55.6%; Score 5; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 6
I40870
phospholipase C (EC 3.1.4.3) - Clostridium perfringens (fragment)
C;Species: Clostridium perfringens
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40870
R;Toyonaga, T.; Matsushita, O.; Katayama, S.; Minami, J.; Okabe, A.
Microbiol. Immunol. 36, 603-613, 1992
A;Title: Role of the upstream region containing an intrinsic DNA curvature in the negati
A;Reference number: I40870; MUID:92396045; PMID:1522810
A;Accession: I40870
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4 <RES>
A;Cross-references: UNIPARC:UPI000011E99B; EMBL:X62825; NID:g40622; PIDN:CAA44636.1; PID
C;Genetics:
A;Gene: plc
C;Keywords: phosphoric diester hydrolase

Query Match 55.6%; Score 5; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 7
A25844
autho-RF amide neuropeptide - sea pansy (Renilla koellikeri)
C;Species: Renilla koellikeri (Koelliker's sea pansy)
C;Date: 21-May-1988 #sequence_revision 30-Sep-1993 #text_change 11-Jul-1997
C;Accession: A25844
R;Grimmelikhuijzen, C.J.P.; Groeger, A.
FEBS Lett. 211, 105-108, 1987
A;Title: Isolation of the neuropeptide pGlu-Gly-Arg-Phe-amide from the pennatulid Renill
A;Reference number: A25844
A;Accession: A25844
A;Molecule type: protein
A;Residues: 1-4 <GRI>
A;Cross-references: UNIPARC:UPI00001733AF
C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;4/Modified site: amidated carboxyl end (Phe) #status experimental

Query Match 55.6%; Score 5; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 8
I61883
protamine P1 - orangutan (fragment)
C;Species: Pongo pygmaeus (orangutan)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C;Accession: I61883
R;Queralt, R.; Oliva, R.
Gene 133, 197-204, 1993
A;Title: Identification of conserved potential regulatory sequences of the protamine-en
A;Reference number: I37013; MUID:94040810; PMID:8224908
A;Accession: I61883
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4 <RES>
A;Cross-references: UNIPARC:UPI000011E9E7; EMBL:Z12146; NID:g38156; PIDN:CAA78130.1; PI

Query Match 55.6%; Score 5; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 9
S47552
ubiquitin - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-May-1995 #sequence_revision 21-Jul-1995 #text_change 17-Mar-1999
C;Accession: S47552
R;Hubbard, M.J.; Carne, A.
Biochim. Biophys. Acta 1200, 191-196, 1994
A;Title: Differential feeding-related regulation of ubiquitin and calbindin(9kDa) in ra
A;Reference number: S47552; MUID:94304928; PMID:8031840
A;Accession: S47552
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-4 <HUB>
A;Cross-references: UNIPARC:UPI000015207C

Query Match 55.6%; Score 5; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 10
I37013
protamine P1 - Cercopithecus patas (fragment)
C;Species: Cercopithecus patas
C;Date: 19-Mar-1997 #sequence_revision 07-Nov-1997 #text_change 21-Jul-2000
C;Accession: I37013
R;Queralt, R.; Oliva, R.
Gene 133, 197-204, 1993
A;Title: Identification of conserved potential regulatory sequences of the protamine-enc
A;Reference number: I37013; MUID:94040810; PMID:8224908
A;Accession: I37013
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4 <RES>
A;Cross-references: UNIPARC:UPI000011E9E7; EMBL:Z12150; NID:g22814; PIDN:CAA78134.1; PID

Query Match 55.6%; Score 5; DB 2; Length 4;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|

```
Db          3 R 3


RESULT 11
I84439
protamine P1 - savannah baboon (fragment)
C;Species: Papio hamadryas doguera (savannah baboon)
C;Date: 19-Mar-1997 #sequence_revision 07-Nov-1997 #text_change 21-Jul-2000
C;Accession: I84439
R;Queralt, R.; Oliva, R.
Gene 133, 197-204, 1993
A;Title: Identification of conserved potential regulatory sequences of the protamine-enc
A;Reference number: I37013; MUID:94040810; PMID:8224908
A;Accession: I84439
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4 <RES>
A;Cross-references: UNIPARC:UPI000011E9E7; EMBL:Z12147; NID:g38134; PIDN:CAA78131.1; PID

  Query Match             55.6%;  Score 5;  DB 2;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 12
A35779
neuropeptide Antho-RNamide - sea anemone (Anthopleura elegantissima)
C;Species: Anthopleura elegantissima
C;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 09-Jul-2004
C;Accession: A35779
R;Grimmelikhuijzen, C.J.P.; Rinehart, K.L.; Jacob, E.; Graff, D.; Reinscheid, R.K.; Noth
Proc. Natl. Acad. Sci. U.S.A. 87, 5410-5414, 1990
A;Title: Isolation of L-3-phenyllactyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea anemone ne
A;Reference number: A35779; MUID:90319122; PMID:1973541
A;Accession: A35779
A;Molecule type: protein
A;Residues: 1-4 <GRI>
A;Cross-references: UNIPROT:P58707; UNIPARC:UPI000012AA36
C;Comment: The L-3-phenyllactyl blocking group probably arises from an amino-terminal ph
C;Keywords: amidated carboxyl end; neuropeptide; phenyllactylation
F;1/Modified site: L-3-phenyllactic acid (Phe) #status experimental
F;4/Modified site: amidated carboxyl end (Asn) #status experimental

  Query Match             55.6%;  Score 5;  DB 2;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 13
A60418
FMRFamide - polychaete (Nereis virens)
C;Species: Nereis virens (sandworm)
C;Date: 11-Feb-1993 #sequence_revision 11-Feb-1993 #text_change 09-Jul-2004
C;Accession: A60418
R;Krajniak, K.G.; Price, D.A.
Peptides 11, 75-77, 1990
A;Title: Authentic FMRFamide is present in the polychaete Nereis virens.
A;Reference number: A60418; MUID:90259866; PMID:2342992
A;Accession: A60418
A;Molecule type: protein
A;Residues: 1-4 <KRA>
A;Cross-references: UNIPROT:P01162; UNIPARC:UPI000012AAD5
C;Keywords: amidated carboxyl end; neuropeptide
F;4/Modified site: amidated carboxyl end (Phe) #status experimental

  Query Match             55.6%;  Score 5;  DB 2;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 14
PT0721
T-cell receptor beta chain V-D-J region (140-2K) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C;Accession: PT0721
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0721
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-4 <FEE>
A;Cross-references: UNIPROT:Q8JZN5; UNIPROT:P54729; UNIPROT:Q8CBY1; UNIPROT:Q8C0C0; UNIP
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

  Query Match             55.6%;  Score 5;  DB 2;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 15
ECNK
cardioexcitatory neuropeptide FMRFamide - sunray clam
C;Species: Macrocallista nimbosa (sunray clam)
C;Date: 20-Jun-2000 #sequence_revision 20-Jun-2000 #text_change 16-Aug-2004
C;Accession: A01426
R;Price, D.A.; Greenberg, M.J.
Science 197, 670-671, 1977
A;Title: Structure of a molluscan cardioexcitatory neuropeptide.
A;Reference number: A01426; MUID:77215956; PMID:877582
A;Accession: A01426
A;Molecule type: protein
A;Residues: 1-4 <PRI>
A;Cross-references: UNIPROT:P01162; UNIPARC:UPI000012AAD5
A;Note: the active peptide was also synthesized
C;Comment: This peptide was purified from pooled extracts of cerebral, pedal, and visce
action in molluscs; its exact physiological role is not yet established.
C;Keywords: amidated carboxyl end; neuropeptide
F;4/Modified site: amidated carboxyl end (Phe) #status experimental

  Query Match             55.6%;  Score 5;  DB 2;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 16
HOROHA
proctolin - American cockroach
C;Species: Periplaneta americana (American cockroach)
C;Date: 29-Jul-1981 #sequence_revision 29-Jul-1981 #text_change 09-Jul-2004
C;Accession: A01644
R;Starratt, A.N.; Brown, B.E.
Life Sci. 17, 1253-1256, 1975
```

A;Title: Structure of the pentapeptide proctolin, a proposed neurotransmitter in insects
A;Reference number: A93048; MUID:76074708; PMID:576
A;Accession: A01644
A;Molecule type: protein
A;Residues: 1-5 <STA>
A;Cross-references: UNIPROT:P01373; UNIPARC:UPI0000132177
A;Note: the synthetic peptide had the same chromatographic, electrophoretic, and pharmac
R;O'Shea, M.; Adams, M.E.
Science 213, 567-569, 1981
A;Title: Pentapeptide (proctolin) associated with an identified neuron.
A;Reference number: A94260; MUID:81225865; PMID:6113690
A;Contents: annotation; biological source
C;Comment: This peptide is found in the lateral white neurons, which occur (in the cockr
innervate the striated hindgut muscles in insects and stimulate contraction of these mus
C;Superfamily: proctolin
C;Keywords: neuropeptide

  Query Match 55.6%; Score 5; DB 1; Length 5;
  Best Local Similarity 100.0%; Pred. No. 0;
  Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
  |
Db 1 R 1

RESULT 17
JN0862
peptidyl-dipeptidase A inhibitory peptide C112 - striped bonito
C;Species: Sarda orientalis (striped bonito)
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 07-May-1999
C;Accession: JN0862
R;Matsumura, N.; Fujii, M.; Takeda, Y.; Shimizu, T.
Biosci. Biotechnol. Biochem. 57, 1743-1744, 1993
A;Title: Isolation and characterization of angiotensin I-converting enzyme inhibitory pe
A;Reference number: JN0859; MUID:94080036; PMID:7764272
A;Accession: JN0862
A;Molecule type: protein
A;Residues: 1-5 <MAT>
A;Cross-references: UNIPARC:UPI00001567C3
A;Experimental source: intestine
C;Comment: The amino terminal tripeptide of this protein inhibits angiotensin I-converti
C;Superfamily: bradykinin-potentiating peptide
C;Keywords: angiotensin-converting enzyme inhibitor

  Query Match 55.6%; Score 5; DB 2; Length 5;
  Best Local Similarity 100.0%; Pred. No. 0;
  Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
  |
Db 2 R 2

RESULT 18
I40702
primase - Citrobacter diversus (fragment)
C;Species: Citrobacter diversus
C;Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 16-Aug-1996
C;Accession: I40702
R;Versalovic, J.; Lupski, J.R.
Mol. Microbiol. 8, 343-355, 1993
A;Title: Conservation and evolution of the rpsU-dnaG-rpoD macromolecular synthesis (MMS)
A;Reference number: I40702; MUID:93302510; PMID:8316085
A;Accession: I40702
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: UNIPARC:UPI000011EC67; GB:L01754; NID:g144439
C;Genetics:
A;Gene: dnaG

  Query Match 55.6%; Score 5; DB 2; Length 5;
  Best Local Similarity 100.0%; Pred. No. 0;
  Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
  |
Db 4 R 4

RESULT 19
A44955
alkanal monooxygenase (FMN-linked) (EC 1.14.14.3) alpha chain - Vibrio harveyi (fragment
C;Species: Vibrio harveyi
C;Date: 03-Jun-1993 #sequence_revision 03-Jun-1993 #text_change 26-May-2000
C;Accession: A44955
R;Paquatte, O.; Tu, S.C.
Photochem. Photobiol. 50, 817-825, 1989
A;Title: Chemical modification and characterization of the alpha cysteine 106 at the Vi
A;Reference number: A44955; MUID:90175700; PMID:2626493
A;Accession: A44955
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-5 <PAQ>
A;Cross-references: UNIPARC:UPI000017AAD8
C;Keywords: FMN; luminescence; monooxygenase; oxidoreductase

  Query Match 55.6%; Score 5; DB 2; Length 5;
  Best Local Similarity 100.0%; Pred. No. 0;
  Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
  |
Db 5 R 5

RESULT 20
I39964
ribosomal protein S4 - Bacillus circulans (fragment)
C;Species: Bacillus circulans
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 19-Jul-1996
C;Accession: I39964
R;Grundy, F.J.; Henkin, T.M.
J. Bacteriol. 174, 6763-6770, 1992
A;Title: Characterization of the Bacillus subtilis rpsD regulatory target site.
A;Reference number: I39963; MUID:93015735; PMID:1400226
A;Accession: I39964
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: UNIPARC:UPI000011EC5E; GB:M99041; NID:g143471
C;Genetics:
A;Gene: rpsD

  Query Match 55.6%; Score 5; DB 2; Length 5;
  Best Local Similarity 100.0%; Pred. No. 0;
  Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
  |
Db 3 R 3

RESULT 21
I39966
ribosomal protein S4 - Bacillus licheniformis (fragment)
C;Species: Bacillus licheniformis
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 19-Jul-1996
C;Accession: I39966
R;Grundy, F.J.; Henkin, T.M.
J. Bacteriol. 174, 6763-6770, 1992
A;Title: Characterization of the Bacillus subtilis rpsD regulatory target site.
A;Reference number: I39963; MUID:93015735; PMID:1400226
A;Accession: I39966
A;Status: preliminary; translated from GB/EMBL/DDBJ

A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: UNIPARC:UPI000011EC5E; GB:M99043; NID:g143475
C;Genetics:
A;Gene: rpsD

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 22
I39965
ribosomal protein S4 - Bacillus megaterium (fragment)
C;Species: Bacillus megaterium
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 19-Jul-1996
C;Accession: I39965
R;Grundy, F.J.; Henkin, T.M.
J. Bacteriol. 174, 6763-6770, 1992
A;Title: Characterization of the Bacillus subtilis rpsD regulatory target site.
A;Reference number: I39963; MUID:93015735; PMID:1400226
A;Accession: I39965
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: UNIPARC:UPI000011EC5E; GB:M99042; NID:g143473
C;Genetics:
A;Gene: rpsD

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 23
D60274
major protein antigen MPT46 - Mycobacterium tuberculosis (fragment)
C;Species: Mycobacterium tuberculosis
C;Date: 11-Dec-1992 #sequence_revision 11-Dec-1992 #text_change 30-Sep-1993
C;Accession: D60274
R;Nagai, S.; Wiker, H.G.; Harboe, M.; Kinomoto, M.
Infect. Immun. 59, 372-382, 1991
A;Title: Isolation and partial characterization of major protein antigens in the culture
A;Reference number: A60274; MUID:91099989; PMID:1898899
A;Accession: D60274
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-5 <NAG>
A;Cross-references: UNIPARC:UPI0000150742

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 1 R 1

RESULT 24
B22565
R-phycoerythrin alpha-2 chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: B22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601; PMID:3886644
A;Accession: B22565
A;Molecule type: protein
A;Residues: 1-5 <KLO>
A;Cross-references: UNIPARC:UPI000017AEC3

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 5 R 5

RESULT 25
F22565
R-phycoerythrin gamma-A chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: F22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601; PMID:3886644
A;Accession: F22565
A;Molecule type: protein
A;Residues: 1-5 <KLO>
A;Cross-references: UNIPARC:UPI000017AEC4

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 5 R 5

RESULT 26
T14910
hypothetical protein - parsley
C;Species: Petroselinum crispum (parsley)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14910
R;Kircher, S.; Ledger, S.; Hayashi, H.; Weisshaar, B.; Schafer, E.; Frohnmeyer, H.
Mol. Gen. Genet. 257, 595-605, 1998
A;Title: CPRF4a, a novel plant bZIP protein of the CPRF family: comparative analysis of
A;Reference number: Z18261; MUID:98265918; PMID:9604882
A;Accession: T14910
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-5 <KIR>
A;Cross-references: UNIPARC:UPI000011E9D7; EMBL:Y10810; NID:g3336904; PIDN:CAA71769.1;
A;Experimental source: ssp. Hamburger Schnitt

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

RESULT 27
A60803
neuropeptide - sea anemone (Anthopleura elegantissima)
C;Species: Anthopleura elegantissima
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999

C;Accession: A60803
R;Graff, D.; Grimmelikhuijzen, C.J.P.
Brain Res. 442, 354-358, 1988
A;Title: Isolation of <Glu-Ser-Lu-Arg-Trp-NH-2, a novel neuropeptide from sea anemones.
A;Reference number: A60803; MUID:88222764; PMID:2897223
A;Accession: A60803
A;Molecule type: protein
A;Residues: 1-5 <GRA>
A;Cross-references: UNIPARC:UPI000017B683
C;Keywords: amidated carboxyl end; neuropeptide; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;5/Modified site: amidated carboxyl end (Trp) #status experimental

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

RESULT 28
S53595
hypothetical protein (upstream of transcription factor, CCAAT-binding) - chicken
C;Species: Gallus gallus (chicken)
C;Date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 07-May-1999
C;Accession: S53595
R;Calkhoven, C.F.; Bouwman, P.R.J.; Snippe, L.; Ab, G.
Nucleic Acids Res. 22, 5540-5547, 1994
A;Title: Translation start site multiplicity of the CCAAT/enhancer binding protein alpha
A;Reference number: S53595; MUID:95140613; PMID:7838705
A;Accession: S53595
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-5 <CAL>
A;Cross-references: UNIPARC:UPI000017C010; EMBL:X66844

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

RESULT 29
PT0295
Ig heavy chain CRD3 region (clone 5-91) - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 16-Aug-1996
C;Accession: PT0295
R;Yamada, M.; Wasserman, R.; Reichard, B.A.; Shane, S.; Caton, A.J.; Rovera, G.
J. Exp. Med. 173, 395-407, 1991
A;Title: Preferential utilization of specific immunoglobulin heavy chain diversity and j
A;Reference number: PT0222; MUID:91108337; PMID:1899102
A;Accession: PT0295
A;Molecule type: DNA
A;Residues: 1-5 <YAM>
A;Cross-references: UNIPARC:UPI000017C20E
A;Experimental source: B lymphocyte
C;Keywords: heterotetramer; immunoglobulin

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

RESULT 30
S62883
seminal plasma protein II - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 17-Mar-1999
C;Accession: S62883
R;Romero, A.; Varela, P.F.; Sanz, L.; Toepfer-Petersen, E.; Calvete, J.J.
FEBS Lett. 382, 15-17, 1996
A;Title: Crystallization and preliminary X-ray diffraction analysis of boar seminal pla
A;Reference number: S62882; MUID:96196555; PMID:8612739
A;Accession: S62883
A;Molecule type: protein
A;Residues: 1-5 <ROM>
A;Cross-references: UNIPARC:UPI000017C475
C;Complex: heterodimer; seminal plasma protein I and seminal plasma protein II
C;Keywords: glycoprotein; heterodimer; semen

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

RESULT 31
PT0513
T-cell receptor beta chain V-D-J region (100-4AL) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0513; PT0606
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0513
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Cross-references: UNIPARC:UPI000017C7BC
A;Experimental source: adult thymus, strain BALB/c, clone 100-4AL
A;Accession: PT0606
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FE2>
A;Cross-references: UNIPARC:UPI000017C7BC
A;Experimental source: newborn thymus, strain BALB/c, clone 120-1S
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 5 R 5

RESULT 32
PT0525
T-cell receptor beta chain V-D-J region (100-4J) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0525
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0525
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>

A;Cross-references: UNIPARC:UPI000017C7C4
A;Experimental source: adult thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 33
PT0597
T-cell receptor beta chain V-D-J region (111-1B) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0597
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0597
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Cross-references: UNIPARC:UPI000017C7CF
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 5 R 5

RESULT 34
PT0672
T-cell receptor beta chain V-D-J region (121-1B) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0646; PT0672
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0646
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Cross-references: UNIPARC:UPI000017C7E9
A;Experimental source: day 4 postnatal thymus, strain BALB/c, clone 121-1B
A;Accession: PT0672
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-5 <FE2>
A;Cross-references: UNIPARC:UPI000017C7E9
A;Experimental source: day 18 fetal thymus, strain BALB/c, clone 140-1BG
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 5 R 5

RESULT 35
PT0553
T-cell receptor beta chain V-D-J region (126-1C) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0553
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0553
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Cross-references: UNIPARC:UPI000017C810
A;Experimental source: day 18 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

RESULT 36
PT0695
T-cell receptor beta chain V-D-J region (135-1D) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0695
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0695
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-5 <FEE>
A;Cross-references: UNIPARC:UPI000017C81B
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

RESULT 37
PT0577
T-cell receptor beta chain V-D-J region (141-1BC) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0577; PT0574
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0577
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Cross-references: UNIPARC:UPI000017C830
A;Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1BC
A;Accession: PT0574
A;Status: translation not shown
A;Molecule type: mRNA

A;Residues: 1-5 <FE2>
A;Cross-references: UNIPARC:UPI000017C830
A;Experimental source: day 19 fetal thymus, strain BALB/c, clone 141-1Q
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

RESULT 38
PT0572
T-cell receptor beta chain V-D-J region (141-1CO) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0572
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0572
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Cross-references: UNIPARC:UPI000017C838
A;Experimental source: day 19 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 5 R 5

RESULT 39
PT0700
T-cell receptor beta chain V-D-J region (161-2A) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 30-May-1997
C;Accession: PT0700
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0700
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-5 <FEE>
A;Cross-references: UNIPARC:UPI000017C841
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

RESULT 40
A60411
proctolin - Atlantic horseshoe crab
C;Species: Limulus polyphemus (Atlantic horseshoe crab)
C;Date: 03-Feb-1993 #sequence_revision 03-Feb-1993 #text_change 09-Jul-2004
C;Accession: A60411
R;Groome, J.R.; Tillinghast, E.K.; Townley, M.A.; Vetrovs, A.; Watson III, W.H.; Hunt,
Peptides 11, 205-211, 1990
A;Title: Identification of proctolin in the central nervous system of the horseshoe cra
A;Reference number: A60411; MUID:90287800; PMID:2356151
A;Accession: A60411
A;Molecule type: protein
A;Residues: 1-5 <GRO>
A;Cross-references: UNIPROT:P01373; UNIPARC:UPI0000132177
C;Comment: This neuropeptide stimulates cardiac output and hindgut motility in the hors
C;Keywords: neuropeptide

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 1 R 1

RESULT 41
PT0608
T-cell receptor beta chain V-D-J region (120-2CF) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C;Accession: PT0608
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0608
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Cross-references: UNIPROT:O18345; UNIPARC:UPI000017C7E0
A;Experimental source: newborn thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

RESULT 42
PT0565
T-cell receptor beta chain V-D-J region (141-1CF) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 09-Jul-2004
C;Accession: PT0565
R;Feeney, A.J.
J. Exp. Med. 174, 115-124, 1991
A;Title: Junctional sequences of fetal T cell receptor beta chains have few N regions.
A;Reference number: PT0509; MUID:91277601; PMID:1711558
A;Accession: PT0565
A;Status: translation not shown
A;Molecule type: mRNA
A;Residues: 1-5 <FEE>
A;Cross-references: UNIPROT:Q8BZQ7; UNIPARC:UPI000017C835
A;Experimental source: day 19 fetal thymus, strain BALB/c
C;Keywords: T-cell receptor

Query Match 55.6%; Score 5; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

```
RESULT 43
A35890
RNA-directed DNA polymerase (EC 2.7.7.49) 66K chain - human immunodeficiency virus type
C;Species: human immunodeficiency virus type 1, HIV-1
C;Date: 09-Nov-1990 #sequence_revision 09-Nov-1990 #text_change 31-Dec-1993
C;Accession: A35890
R;Bathurst, I.C.; Moen, L.K.; Lujan, M.A.; Gibson, H.L.; Feucht, P.H.; Pichuantes, S.; C
Biochem. Biophys. Res. Commun. 171, 589-595, 1990
A;Title: Characterization of the human immunodeficiency virus type-1 reverse transcripta
A;Reference number: A35890; MUID:90386627; PMID:1698361
A;Accession: A35890
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-6 <BAT>
A;Cross-references: UNIPARC:UPI000017A88C
C;Keywords: nucleotidyltransferase

  Query Match             55.6%;  Score 5;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 44
A37765
hypothetical protein (csmA 5' region) - Chloroflexus aurantiacus (fragment)
C;Species: Chloroflexus aurantiacus
C;Date: 31-May-1991 #sequence_revision 31-May-1991 #text_change 30-Sep-1993
C;Accession: A37765
R;Theroux, S.J.; Redlinger, T.E.; Fuller, R.C.; Robinson, S.J.
J. Bacteriol. 172, 4497-4504, 1990
A;Title: Gene encoding the 5.7-kilodalton chlorosome protein of Chloroflexus aurantiacus
A;Reference number: A37765; MUID:90330558; PMID:2376566
A;Accession: A37765
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-6 <THE>
A;Cross-references: UNIPARC:UPI000017ABA1; GB:M33964

  Query Match             55.6%;  Score 5;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          6 R 6


RESULT 45
C22565
R-phycoerythrin beta-1 chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: C22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601; PMID:3886644
A;Accession: C22565
A;Molecule type: protein
A;Residues: 1-6 <KLO>
A;Cross-references: UNIPARC:UPI0000171E4A

  Query Match             55.6%;  Score 5;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          6 R 6


RESULT 46
PQ0008
angiotensin-converting enzyme inhibitor (FLP-1) - common fig
N;Alternate names: ficus latex peptide 1
C;Species: Ficus carica (common fig)
C;Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 08-Dec-1995
C;Accession: PQ0008
R;Maruyama, S.; Miyoshi, S.; Tanaka, H.
Agric. Biol. Chem. 53, 2763-2767, 1989
A;Title: Angiotensin I-converting enzyme inhibitors derived from Ficus carica.
A;Reference number: PQ0008
A;Accession: PQ0008
A;Molecule type: protein
A;Residues: 1-6 <MAR>
A;Cross-references: UNIPARC:UPI000015655D
A;Experimental source: latex
C;Keywords: angiotensin-converting enzyme inhibitor

  Query Match             55.6%;  Score 5;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          6 R 6


RESULT 47
A60494
antineoplastic glycoprotein - sea hare (Dolabella auricularia) (fragment)
N;Alternate names: dolabellanin C
C;Species: Dolabella auricularia
C;Date: 19-Mar-1993 #sequence_revision 19-Mar-1993 #text_change 18-Jun-1993
C;Accession: A60494
R;Kisugi, J.; Kamiya, H.; Yamazaki, M.
Dev. Comp. Immunol. 13, 3-8, 1989
A;Title: Purification of dolabellanin-C an antineoplastic glycoprotein in the body flui
A;Reference number: A60494; MUID:89357188; PMID:2767307
A;Accession: A60494
A;Molecule type: protein
A;Residues: 1-6 <KIS>
A;Cross-references: UNIPARC:UPI000017BD9D
C;Keywords: cytolysis; glycoprotein; trimer

  Query Match             55.6%;  Score 5;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          5 R 5


RESULT 48
I51434
H4 histone - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: I51434
R;Woodland, H.R.; Warmington, J.R.; Ballantine, J.E.M.; Turner, P.C.
Nucleic Acids Res. 12, 4939-4958, 1984
A;Title: Are there major developmentally regulated H4 gene classes in Xenopus?.
A;Reference number: I51391; MUID:84247348; PMID:6330691
A;Accession: I51434
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-6 <WOO>
A;Cross-references: UNIPARC:UPI000011E896; GB:K02304; NID:g214227; PIDN:AAA49738.1; PID
```

Query Match 55.6%; Score 5; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 4 R 4

RESULT 49
I37027
protamine P1 - gorilla (fragment)
C;Species: Gorilla gorilla (gorilla)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C;Accession: I37027
R;Queralt, R.; Oliva, R.
Gene 133, 197-204, 1993
A;Title: Identification of conserved potential regulatory sequences of the protamine-enc
A;Reference number: I37013; MUID:94040810; PMID:8224908
A;Accession: I37027
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-6 <RES>
A;Cross-references: UNIPARC:UPI000011E9E8; EMBL:Z12145; NID:g22910; PIDN:CAA78129.1; PID

Query Match 55.6%; Score 5; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 3 R 3

RESULT 50
A11490
pyruvate kinase (EC 2.7.1.40) - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 03-Mar-1995
C;Accession: A11490
R;Hjelmquist, G.; Andersson, J.; Edlund, B.; Engstrom, L.
Biochem. Biophys. Res. Commun. 61, 559-563, 1974
A;Title: Amino acid sequence of a (32-P)phosphopeptide from pig liver pyruvate kinase ph
A;Reference number: A11490; MUID:75127438; PMID:4375989
A;Accession: A11490
A;Molecule type: protein
A;Residues: 1-6 <HJE>
A;Cross-references: UNIPARC:UPI000017C474
A;Experimental source: liver
C;Keywords: glycolysis; phosphotransferase

Query Match 55.6%; Score 5; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 R 1
|
Db 2 R 2

Search completed: January 25, 2006, 18:41:27
Job time : 18.5 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:33:23 ; Search time 63 Seconds
(without alignments)
55.994 Million cell updates/sec

Title: US-10-771-242-293
Perfect score: 9
Sequence: 1 RXXXX 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	5	55.6	4	1	FAR3_HIRME	P42562 hirudo medi
2	5	55.6	4	1	FAR4_HIRME	P42563 hirudo medi
3	5	55.6	4	1	FLRF_HELTI	P69138 helisoma tr
4	5	55.6	4	1	FLRF_HIRME	P69137 hirudo medi
5	5	55.6	4	1	FLRN_ANTEL	P58707 anthopleura
6	5	55.6	4	1	FMRF_HELTI	P69148 helisoma tr
7	5	55.6	4	1	FMRF_HIRME	P69147 hirudo medi
8	5	55.6	4	1	FMRF_MACNI	P69145 macrocallis
9	5	55.6	4	1	FMRF_NERVI	P69146 nereis vire
10	5	55.6	4	1	FYRI_ANTEL	P58706 anthopleura
11	5	55.6	4	1	TUFT_HUMAN	P01858 homo sapien
12	5	55.6	5	1	FARP_ARTTR	P41853 artioposthi
13	5	55.6	5	1	FARP_CHICK	P83308 gallus gall
14	5	55.6	5	1	PRCT_CARMA	P67857 carcinus ma
15	5	55.6	5	1	PRCT_LIMPO	P67858 limulus pol
16	5	55.6	5	1	PRCT_PERAM	P67859 periplaneta
17	5	55.6	5	1	UF01_MOUSE	P38639 mus musculu
18	5	55.6	6	1	ACPH_RABIT	P25154 oryctolagus
19	5	55.6	6	1	FARP_MONEX	P41966 moniezia ex
20	5	55.6	6	1	PYF1_PENMO	P84005 penaeus mon
21	5	55.6	6	2	P82181_SPIOL	P82181 spinacia ol
22	5	55.6	6	2	P82182_SPIOL	P82182 spinacia ol
23	5	55.6	6	2	P82541_SPIOL	P82541 spinacia ol
24	5	55.6	7	1	CAP6_CANAL	P83784 candida alb
25	5	55.6	7	1	CARP_MYTED	P10420 mytilus edu
26	5	55.6	7	1	CHOX_ALCSP	P16101 alcaligenes
27	5	55.6	7	1	FAF1_ASCSU	P31889 ascaris suu
28	5	55.6	7	1	FAF2_ASCSU	P67879 ascaris suu
29	5	55.6	7	1	FAF2_PANRE	P67880 panagrellus
30	5	55.6	7	1	FAR1_HELTI	P41871 helisoma tr
31	5	55.6	7	1	FAR1_MACRS	P83274 macrobrachi
32	5	55.6	7	1	FAR1_PROCL	P38499 procambarus
33	5	55.6	7	1	FAR2_PROCL	P38498 procambarus
34	5	55.6	7	1	FAR3_HAECO	P81298 haemonchus
35	5	55.6	7	1	FAR3_PANRE	P41874 panagrellus
36	5	55.6	7	1	FAR4_PANRE	P41875 panagrellus
37	5	55.6	7	1	FAR5_HIRME	P42564 hirudo medi
38	5	55.6	7	1	FARB_CALVO	P41866 calliphora
39	5	55.6	7	1	HCYB_CONCC	P84620 concholepas
40	5	55.6	7	1	IPYR_CANAL	P83777 candida alb
41	5	55.6	7	1	LANC_CARUI	P36960 carnobacter
42	5	55.6	7	1	TY51_LITRU	P82065 litoria rub
43	5	55.6	7	1	UF04_MOUSE	P38642 mus musculu
44	5	55.6	7	1	UH11_RAT	P56576 rattus norv
45	5	55.6	7	1	WWA1_ACHFU	P35919 achatina fu
46	5	55.6	7	2	Q95945_YEAST	Q95945 saccharomyc
47	5	55.6	7	2	Q15903_HUMAN	Q15903 homo sapien
48	5	55.6	7	2	Q8NHH7_HUMAN	Q8nhh7 homo sapien
49	5	55.6	7	2	Q8TAQ4_HUMAN	Q8taq4 homo sapien
50	5	55.6	7	2	O98866_SPIOL	O98866 spinacia ol
51	5	55.6	7	2	P92210_AGRCR	P92210 agropyron c
52	5	55.6	7	2	P92214_9POAL	P92214 amblyopyrum
53	5	55.6	7	2	P92218_9POAL	P92218 australopyr
54	5	55.6	7	2	P92221_BROIN	P92221 bromus iner
55	5	55.6	7	2	P92226_CRIDE	P92226 crithopsis
56	5	55.6	7	2	P92372_9POAL	P92372 haynaldia v
57	5	55.6	7	2	P92381_9POAL	P92381 hordeum bra
58	5	55.6	7	2	P92385_HORMA	P92385 hordeum mar
59	5	55.6	7	2	P92387_9POAL	P92387 henrardia p
60	5	55.6	7	2	P92390_HETPI	P92390 heteranthel
61	5	55.6	7	2	P92393_HORVU	P92393 hordeum vul
62	5	55.6	7	2	P92403_LOPEL	P92403 lophopyrum
63	5	55.6	7	2	P92421_PSAFR	P92421 psathyrosta
64	5	55.6	7	2	P92425_PSEPI	P92425 pseudoroegn
65	5	55.6	7	2	P92427_9POAL	P92427 peridictyon
66	5	55.6	7	2	P92430_AEGTA	P92430 aegilops ta
67	5	55.6	7	2	P92440_THIBE	P92440 thinopyrum
68	5	55.6	7	2	P92442_TAECM	P92442 taeniatheru
69	5	55.6	7	2	P93233_LYCES	P93233 lycopersico
70	5	55.6	7	2	P84495_CUCMA	P84495 cucurbita m
71	5	55.6	7	2	O34028_9SPHN	O34028 sphingomona
72	5	55.6	7	2	O50556_ACTAC	O50556 actinobacil
73	5	55.6	7	2	Q47477_ECOLI	Q47477 escherichia
74	5	55.6	7	2	O07354_SYNP8	O07354 synechococc
75	5	55.6	7	2	Q47505_ECOLI	Q47505 escherichia
76	5	55.6	7	2	Q63480_RAT	Q63480 rattus norv
77	5	55.6	7	2	Q66113_9COMO	Q66113 cherry leaf
78	5	55.6	7	2	Q67113_9INFA	Q67113 influenza a
79	5	55.6	7	2	Q9YIQ9_ADE04	Q9yiq9 human adeno
80	5	55.6	7	2	Q9YIR0_ADE07	Q9yir0 human adeno
81	5	55.6	7	2	Q9YVE3_ADE07	Q9yve3 human adeno
82	5	55.6	8	1	ACT_CARMA	P80709 carcinus ma
83	5	55.6	8	1	ALL3_CYDPO	P82154 cydia pomon
84	5	55.6	8	1	ALL4_CALVO	P41840 calliphora
85	5	55.6	8	1	ALL4_CYDPO	P82155 cydia pomon
86	5	55.6	8	1	ALL5_CYDPO	P82156 cydia pomon
87	5	55.6	8	1	ANGT2_BOTJA	Q10582 bothrops ja
88	5	55.6	8	1	B44K_FORGI	P81886 porphyromon
89	5	55.6	8	1	C125_CYPDO	P83661 cyphononyx
90	5	55.6	8	1	COX6B_RAT	P80430 rattus norv
91	5	55.6	8	1	FAR1_PANRE	P41872 panagrellus
92	5	55.6	8	1	FAR1_PENMO	P83316 penaeus mon
93	5	55.6	8	1	FAR2_MACRS	P83275 macrobrachi
94	5	55.6	8	1	FAR3_HOMAM	P41486 homarus ame
95	5	55.6	8	1	FAR4_HOMAM	P41487 homarus ame
96	5	55.6	8	1	FAR4_MACRS	P83277 macrobrachi
97	5	55.6	8	1	FAR7_ASCSU	P43171 ascaris suu
98	5	55.6	8	1	FAR8_CALVO	P41863 calliphora
99	5	55.6	8	1	KINI1_PERAM	P82685 periplaneta
100	5	55.6	8	1	LMT2_LOCMI	P22396 locusta mig
101	5	55.6	8	1	LPK_LEUMA	P13049 leucophaea
102	5	55.6	8	1	LPMS_STAEP	P23211 staphylococ
103	5	55.6	8	1	NPMB_BOVIN	P15507 bos taurus
104	5	55.6	8	1	NS3_MYCTU	P81152 mycobacteri

105	5	55.6	8	1	PPK2_PERAM	P82692	periplaneta
106	5	55.6	8	1	PPK3_PERAM	P82618	periplaneta
107	5	55.6	8	1	RS7_MYCIT	P33564	mycobacteri
108	5	55.6	8	1	UC26_MAIZE	P80632	zea mays (m
109	5	55.6	8	1	UH09_RAT	P56575	rattus norv
110	5	55.6	8	1	UPA1_HUMAN	P30087	homo sapien
111	5	55.6	8	2	Q7M4U4_ASPFI	Q7m4u4	aspergillus
112	5	55.6	8	2	Q9T2W0_YEAST	Q9t2w0	saccharomyc
113	5	55.6	8	2	Q15888_HUMAN	Q15888	homo sapien
114	5	55.6	8	2	Q15898_HUMAN	Q15898	homo sapien
115	5	55.6	8	2	Q15900_HUMAN	Q15900	homo sapien
116	5	55.6	8	2	Q16468_HUMAN	Q16468	homo sapien
117	5	55.6	8	2	Q69YH8_HUMAN	Q69yh8	homo sapien
118	5	55.6	8	2	Q71UR9_HUMAN	Q71ur9	homo sapien
119	5	55.6	8	2	Q7Z4N9_HUMAN	Q7z4n9	homo sapien
120	5	55.6	8	2	Q8IVK3_HUMAN	Q8ivk3	homo sapien
121	5	55.6	8	2	Q8TF70_HUMAN	Q8tf70	homo sapien
122	5	55.6	8	2	Q9Y4J4_HUMAN	Q9y4j4	homo sapien
123	5	55.6	8	2	Q53SB0_HUMAN	Q53sb0	homo sapien
124	5	55.6	8	2	Q9UMH9_HUMAN	Q9umh9	homo sapien
125	5	55.6	8	2	Q75MD1_HUMAN	Q75md1	homo sapien
126	5	55.6	8	2	Q4EW64_HUMAN	Q4ew64	homo sapien
127	5	55.6	8	2	O15896_BABBO	O15896	babesia bov
128	5	55.6	8	2	O15899_BABOV	O15899	babesia ovi
129	5	55.6	8	2	Q70MX3_TRYBR	Q70mx3	trypanosoma
130	5	55.6	8	2	Q7M3N2_MANSE	Q7m3n2	manduca sex
131	5	55.6	8	2	Q86BS9_STRPU	Q86bs9	strongyloce
132	5	55.6	8	2	Q94623_MANSE	Q94623	manduca sex
133	5	55.6	8	2	Q9N6M5_TOXGO	Q9n6m5	toxoplasma
134	5	55.6	8	2	Q7GEM6_BRALA	Q7gem6	branchiosto
135	5	55.6	8	2	Q8WGD7_9EUCA	Q8wgd7	lomis hirta
136	5	55.6	8	2	Q4Y0C0_PLACH	Q4y0c0	plasmodium
137	5	55.6	8	2	Q7M4A4_MERMC	Q7m4a4	mercenaria
138	5	55.6	8	2	Q28866_MEGNO	Q28866	megaptera n
139	5	55.6	8	2	Q5RC44_PONPY	Q5rc44	pongo pygma
140	5	55.6	8	2	Q5RLS9_PIG	Q5rls9	sus scrofa
141	5	55.6	8	2	Q9BF92_TURTR	Q9bf92	tursiops tr
142	5	55.6	8	2	Q9BF93_MEGNO	Q9bf93	megaptera n
143	5	55.6	8	2	Q9GMH3_LAGOB	Q9gmh3	lagenorhync
144	5	55.6	8	2	Q70KG9_PIG	Q70kg9	sus scrofa
145	5	55.6	8	2	Q4G3W0_MACMU	Q4g3w0	macaca mula
146	5	55.6	8	2	Q37854_BPR17	Q37854	bacteriopha
147	5	55.6	8	2	Q5W4V4_9VIRU	Q5w4v4	bacteriopha
148	5	55.6	8	2	Q5W4V6_9CAUD	Q5w4v6	bacteriopha
149	5	55.6	8	2	Q5W4V8_BPR51	Q5w4v8	bacteriopha
150	5	55.6	8	2	Q5W4W1_9CAUD	Q5w4w1	bacteriopha
151	5	55.6	8	2	Q5W4W2_BPR32	Q5w4w2	bacteriopha
152	5	55.6	8	2	Q5W4W4_9CAUD	Q5w4w4	bacteriopha
153	5	55.6	8	2	Q5W4W6_9VIRU	Q5w4w6	bacteriopha
154	5	55.6	8	2	Q5W4W8_9CAUD	Q5w4w8	bacteriopha
155	5	55.6	8	2	Q5W4X0_9VIRU	Q5w4x0	bacteriopha
156	5	55.6	8	2	Q5W4X2_9VIRU	Q5w4x2	bacteriopha
157	5	55.6	8	2	Q5W4X4_9VIRU	Q5w4x4	bacteriopha
158	5	55.6	8	2	Q5W4X6_BPLZ5	Q5w4x6	bacteriopha
159	5	55.6	8	2	Q5W4X8_BPLZ4	Q5w4x8	bacteriopha
160	5	55.6	8	2	Q5W4Y0_BPLZ1	Q5w4y0	bacteriopha
161	5	55.6	8	2	Q5W4Y2_9CAUD	Q5w4y2	bacteriopha
162	5	55.6	8	2	Q5W4Y4_9CAUD	Q5w4y4	bacteriopha
163	5	55.6	8	2	Q5W4Y6_BPR70	Q5w4y6	bacteriopha
164	5	55.6	8	2	Q6LDY8_BPF1	Q6ldy8	bacteriopha
165	5	55.6	8	2	Q8H9H3_9CAUD	Q8h9h3	bacteriopha
166	5	55.6	8	2	Q8H9H5_BPT6	Q8h9h5	bacteriopha
167	5	55.6	8	2	Q8H9H6_BPT2	Q8h9h6	bacteriopha
168	5	55.6	8	2	Q8H9I2_BPR27	Q8h9i2	bacteriopha
169	5	55.6	8	2	Q8H9I3_9CAUD	Q8h9i3	bacteriopha
170	5	55.6	8	2	Q8H9I4_9CAUD	Q8h9i4	bacteriopha
171	5	55.6	8	2	Q8H9I6_BPR15	Q8h9i6	bacteriopha
172	5	55.6	8	2	Q8H9I8_9CAUD	Q8h9i8	bacteriopha
173	5	55.6	8	2	Q8H9I9_BPR10	Q8h9i9	bacteriopha
174	5	55.6	8	2	Q8H9J1_BPR06	Q8h9j1	bacteriopha
175	5	55.6	8	2	Q8H9J3_BPR03	Q8h9j3	bacteriopha
176	5	55.6	8	2	Q8H9J5_9CAUD	Q8h9j5	bacteriopha
177	5	55.6	8	2	Q8H9J7_BPOX2	Q8h9j7	bacteriopha
178	5	55.6	8	2	Q8H9J9_BPM1	Q8h9j9	bacteriopha
179	5	55.6	8	2	Q8H9K4_BPK3	Q8h9k4	bacteriopha
180	5	55.6	8	2	O19956_GOSAR	O19956	gossypium a
181	5	55.6	8	2	O19957_GOSHI	O19957	gossypium h
182	5	55.6	8	2	O19958_GOSBA	O19958	gossypium b
183	5	55.6	8	2	O19959_GOSTO	O19959	gossypium t
184	5	55.6	8	2	O19960_GOSMU	O19960	gossypium m
185	5	55.6	8	2	O19961_GOSDA	O19961	gossypium d
186	5	55.6	8	2	Q36898_9SOLA	Q36898	nicotiana p
187	5	55.6	8	2	Q5D4X1_9MYRT	Q5d4x1	physocalymm
188	5	55.6	8	2	Q5D4X2_9MYRT	Q5d4x2	pehria comp
189	5	55.6	8	2	Q5D4X4_9MYRT	Q5d4x4	koehneria m
190	5	55.6	8	2	Q5D4X5_9MYRT	Q5d4x5	nesaea aspe
191	5	55.6	8	2	Q5D4Y0_9MYRT	Q5d4y0	lafoensia a
192	5	55.6	8	2	Q5D4Y1_9MYRT	Q5d4y1	ginoria ame
193	5	55.6	8	2	Q5D4Y5_9MYRT	Q5d4y5	capuronia m
194	5	55.6	8	2	Q5IA47_9MAGN	Q5ia47	laurus nobi
195	5	55.6	8	2	Q5Y9D8_9BRYO	Q5y9d8	warnstorfia
196	5	55.6	8	2	Q659Q3_9CARY	Q659q3	silene oste
197	5	55.6	8	2	Q659Q5_9CARY	Q659q5	silene invo
198	5	55.6	8	2	Q6H0C6_9TRAC	Q6h0c6	isoetes fla
199	5	55.6	8	2	Q6YLT8_SCIVE	Q6ylt8	sciadopitys
200	5	55.6	8	2	Q6ZZ02_9CARY	Q6zz02	lychnis cor
201	5	55.6	8	2	Q70Y68_9LAMI	Q70y68	prostanther
202	5	55.6	8	2	Q70Y69_9LAMI	Q70y69	syncolostem
203	5	55.6	8	2	Q70Y84_9LAMI	Q70y84	plectranthu
204	5	55.6	8	2	Q7GEF5_NICBI	Q7gef5	nicotiana b
205	5	55.6	8	2	Q7IMR4_BRANA	Q7imr4	brassica na
206	5	55.6	8	2	Q7M1F1_RAPSA	Q7m1f1	raphanus sa
207	5	55.6	8	2	Q7M267_GASCO	Q7m267	gastrocloni
208	5	55.6	8	2	Q7XB03_MAIZE	Q7xb03	zea mays (m
209	5	55.6	8	2	Q8GTG5_LYCES	Q8gtg5	lycopersico
210	5	55.6	8	2	Q8L802_MAIZE	Q8l802	zea mays (m
211	5	55.6	8	2	Q8MC14_9MYRT	Q8mc14	fuchsia hyb
212	5	55.6	8	2	Q8MC16_9MYRT	Q8mc16	ludwigia hy
213	5	55.6	8	2	Q8MC18_QUIIN	Q8mc18	quisqualis
214	5	55.6	8	2	Q8MC20_9MYRT	Q8mc20	combretum w
215	5	55.6	8	2	Q8MC22_9MYRT	Q8mc22	sonneratia
216	5	55.6	8	2	Q8MC24_9MYRT	Q8mc24	sonneratia
217	5	55.6	8	2	Q8MC26_9MYRT	Q8mc26	lagerstroem
218	5	55.6	8	2	Q8MC28_9MYRT	Q8mc28	duabanga gr
219	5	55.6	8	2	Q8MC30_LAGSP	Q8mc30	lagerstroem
220	5	55.6	8	2	Q8MC32_9MYRT	Q8mc32	rotala indi
221	5	55.6	8	2	Q8MC34_9MYRT	Q8mc34	heimia myrt
222	5	55.6	8	2	Q8MC40_9MYRT	Q8mc40	nesaea lued
223	5	55.6	8	2	Q8MC42_9MYRT	Q8mc42	sonneratia
224	5	55.6	8	2	Q8MC46_9MYRT	Q8mc46	peplis port
225	5	55.6	8	2	Q8MC48_9MYRT	Q8mc48	pemphis aci
226	5	55.6	8	2	Q8MC50_CUPLA	Q8mc50	cuphea lanc
227	5	55.6	8	2	Q8MC52_9MYRT	Q8mc52	woodfordia
228	5	55.6	8	2	Q8MES7_USEUD	Q8mes7	daphniphyll
229	5	55.6	8	2	Q8MES9_CERJA	Q8mes9	cercidiphyl
230	5	55.6	8	2	Q8MET1_USEUD	Q8met1	saxifraga s
231	5	55.6	8	2	Q9GCZ4_9LILI	Q9gcz4	nenga pumil
232	5	55.6	8	2	Q9GD00_9LILI	Q9gd00	masoala mad
233	5	55.6	8	2	Q9GD47_9LILI	Q9gd47	hydriastele
234	5	55.6	8	2	Q56ZT0_ARATH	Q56zt0	arabidopsis
235	5	55.6	8	2	Q8LW32_9MYRT	Q8lw32	trapa maxim
236	5	55.6	8	2	Q8LW33_PUNGR	Q8lw33	punica gran
237	5	55.6	8	2	Q8LW34_LYTSA	Q8lw34	lythrum sal
238	5	55.6	8	2	Q8MC36_9MYRT	Q8mc36	lawsonia in
239	5	55.6	8	2	Q8MC38_9MYRT	Q8mc38	ammannia ba
240	5	55.6	8	2	Q8MC44_DECVE	Q8mc44	decodon ver
241	5	55.6	8	2	Q4JLA6_BETVU	Q4jla6	beta vulgar
242	5	55.6	8	2	Q6ZZ01_SILCO	Q6zz01	silene coni
243	5	55.6	8	2	O32560_ECOLI	O32560	escherichia
244	5	55.6	8	2	O52062_BACME	O52062	bacillus me
245	5	55.6	8	2	P72279_RHOGO	P72279	rhodococcus
246	5	55.6	8	2	Q45889_CLOBO	Q45889	clostridium
247	5	55.6	8	2	Q51594_9ZZZZ	Q51594	plasmid col
248	5	55.6	8	2	Q56140_STRTR	Q56140	streptococc
249	5	55.6	8	2	Q6DQZ1_CLOBE	Q6dqz1	clostridium
250	5	55.6	8	2	Q6LAA8_SHIFL	Q6laa8	shigella fl

251	5	55.6	8	2	Q6UNM3_VIBCH	Q6unm3	vibrio chol	324	5	55.6	8	2	Q94V88_9SAUR	Q94v88	varanus tri
252	5	55.6	8	2	Q71UF7_MORMO	Q71uf7	morganella	325	5	55.6	8	2	Q94V91_VARTI	Q94v91	varanus tim
253	5	55.6	8	2	Q71V47_PSEPU	Q71v47	pseudomonas	326	5	55.6	8	2	Q94VA0_9SAUR	Q94va0	varanus sem
254	5	55.6	8	2	Q799V9_BACSH	Q799v9	bacillus sp	327	5	55.6	8	2	Q94VA7_9SAUR	Q94va7	varanus sal
255	5	55.6	8	2	Q79CX6_MYXXA	Q79cx6	myxococcus	328	5	55.6	8	2	Q94VB2_9SAUR	Q94vb2	varanus sal
256	5	55.6	8	2	Q7AUJ4_9GAMM	Q7auj4	acinetobact	329	5	55.6	8	2	Q94VB5_9SAUR	Q94vb5	varanus sal
257	5	55.6	8	2	Q7BW19_VIBCH	Q7bw19	vibrio chol	330	5	55.6	8	2	Q94VC1_VARRU	Q94vc1	varanus rud
258	5	55.6	8	2	Q7DKL7_STAAU	Q7dkl7	staphylococ	331	5	55.6	8	2	Q94VE4_VARML	Q94ve4	varanus mel
259	5	55.6	8	2	Q7M194_ECOLI	Q7m194	escherichia	332	5	55.6	8	2	Q94VF3_9SAUR	Q94vf3	varanus kei
260	5	55.6	8	2	Q8RJI0_STRCS	Q8rji0	streptomyce	333	5	55.6	8	2	Q94VF6_VARJO	Q94vf6	varanus job
261	5	55.6	8	2	Q934S4_THIFE	Q934s4	thiobacillu	334	5	55.6	8	2	Q94VF9_VARIN	Q94vf9	varanus ind
262	5	55.6	8	2	Q9R9C2_BORBU	Q9r9c2	borrelia bu	335	5	55.6	8	2	Q94VJ4_VARBN	Q94vj4	varanus ben
263	5	55.6	8	2	Q9S6D5_ECOLI	Q9s6d5	escherichia	336	5	55.6	8	2	Q9PS69_CHICK	Q9ps69	gallus gall
264	5	55.6	8	2	Q9ZIE9_NEIME	Q9zie9	neisseria m	337	5	55.6	8	2	Q9TD02_9SMEG	Q9td02	terranatos
265	5	55.6	8	2	O09258_SYNP8	O09258	synechococc	338	5	55.6	8	2	Q6VMC6_9PASS	Q6vmc6	serilophus
266	5	55.6	8	2	Q79AG6_ENTAG	Q79ag6	enterobacte	339	5	55.6	9	1	BRK1_RANNI	Q71z54	rana nigrom
267	5	55.6	8	2	Q62721_RAT	Q62721	rattus norv	340	5	55.6	9	1	BSP43_SERPL	P83375	serratia pl
268	5	55.6	8	2	Q7M056_9MURI	Q7m056	mus sp. gen	341	5	55.6	9	1	CB22_SPIOL	Q9t2k9	spinacia ol
269	5	55.6	8	2	Q80WD6_MOUSE	Q80wd6	mus musculu	342	5	55.6	9	1	CB2B_SPIOL	Q9t210	spinacia ol
270	5	55.6	8	2	Q80XV8_9MURI	Q80xv8	rattus sp.	343	5	55.6	9	1	CB2C_SPIOL	Q9t211	spinacia ol
271	5	55.6	8	2	Q8CJ03_MOUSE	Q8cj03	mus musculu	344	5	55.6	9	1	CONO_CONGE	P05486	conus geogr
272	5	55.6	8	2	Q8R4D8_MOUSE	Q8r4d8	mus musculu	345	5	55.6	9	1	CONO_CONST	P05487	conus stria
273	5	55.6	8	2	Q8R5M9_MOUSE	Q8r5m9	mus musculu	346	5	55.6	9	1	DNF1_LOCMI	P16339	locusta mig
274	5	55.6	8	2	Q99MH2_MOUSE	Q99mh2	mus musculu	347	5	55.6	9	1	FAR1_CALVO	P41856	calliphora
275	5	55.6	8	2	Q99P40_MOUSE	Q99p40	mus musculu	348	5	55.6	9	1	FAR2_CALVO	P41857	calliphora
276	5	55.6	8	2	Q9ERD2_MOUSE	Q9erd2	mus musculu	349	5	55.6	9	1	FAR2_PANRE	P41873	panagrellus
277	5	55.6	8	2	Q9ET16_MESAU	Q9et16	mesocricetu	350	5	55.6	9	1	FAR3_CALVO	P41858	calliphora
278	5	55.6	8	2	Q9ET17_MUSCR	Q9et17	mus caroli	351	5	55.6	9	1	FAR3_MACRS	P83276	macrobrachi
279	5	55.6	8	2	Q9ET18_MUSSP	Q9et18	mus spretus	352	5	55.6	9	1	FAR3_PENMO	P83318	penaeus mon
280	5	55.6	8	2	Q9JLD7_MESAU	Q9jld7	mesocricetu	353	5	55.6	9	1	FAR4_CALVO	P41859	calliphora
281	5	55.6	8	2	Q9QVJ8_9MURI	Q9qvj8	mus sp. mep	354	5	55.6	9	1	FAR4_PENMO	P83319	penaeus mon
282	5	55.6	8	2	Q78ED1_RAT	Q78ed1	rattus norv	355	5	55.6	9	1	FAR5_ASCSU	P43170	ascaris suu
283	5	55.6	8	2	Q80WD5_MUSSP	Q80wd5	mus spretus	356	5	55.6	9	1	FAR5_CALVO	P41860	calliphora
284	5	55.6	8	2	O89965_POVJC	O89965	polyomaviru	357	5	55.6	9	1	FAR5_PANRE	P82661	panagrellus
285	5	55.6	8	2	Q6PUD5_SV40	Q6pud5	simian viru	358	5	55.6	9	1	FAR5_PENMO	P83320	penaeus mon
286	5	55.6	8	2	Q6PUD7_SV40	Q6pud7	simian viru	359	5	55.6	9	1	FAR6_CALVO	P41861	calliphora
287	5	55.6	8	2	Q6PUD9_SV40	Q6pud9	simian viru	360	5	55.6	9	1	FAR6_MACRS	P83279	macrobrachi
288	5	55.6	8	2	Q6PUE1_SV40	Q6pue1	simian viru	361	5	55.6	9	1	FAR7_CALVO	P41862	calliphora
289	5	55.6	8	2	Q6QSX9_SV40	Q6qsx9	simian viru	362	5	55.6	9	1	FAR8_MACRS	P83281	macrobrachi
290	5	55.6	8	2	Q7T863_9VIRU	Q7t863	largemouth	363	5	55.6	9	1	FAR9_ASCSU	P43172	ascaris suu
291	5	55.6	8	2	Q80H91_9PARA	Q80h91	newcastle d	364	5	55.6	9	1	FARA_CALVO	P41865	calliphora
292	5	55.6	8	2	Q83977_9INFA	Q83977	influenza a	365	5	55.6	9	1	FARD_CALVO	P41868	calliphora
293	5	55.6	8	2	Q84156_9POXV	Q84156	orf virus.	366	5	55.6	9	1	FARP_CALSI	P38495	callinectes
294	5	55.6	8	2	Q84271_HPV19	Q84271	human papil	367	5	55.6	9	1	FIBB_ERYPA	P19346	erythrocebu
295	5	55.6	8	2	Q84273_HPV25	Q84273	human papil	368	5	55.6	9	1	FIBB_MACFU	P19345	macaca fusc
296	5	55.6	8	2	Q9DSN0_9VIRU	Q9dsn0	beet soil-b	369	5	55.6	9	1	FIBB_PAPAN	P19344	papio anubi
297	5	55.6	8	2	Q9DSN1_9VIRU	Q9dsn1	beet soil-b	370	5	55.6	9	1	FIBB_PAPHA	P19343	papio hamad
298	5	55.6	8	2	Q9DSN2_9VIRU	Q9dsn2	beet soil-b	371	5	55.6	9	1	FIBB_THEGE	P19342	theropithec
299	5	55.6	8	2	Q9DSN3_9VIRU	Q9dsn3	beet soil-b	372	5	55.6	9	1	FMRF1_SARBU	P83350	sarcophaga
300	5	55.6	8	2	Q9DSN4_9VIRU	Q9dsn4	beet soil-b	373	5	55.6	9	1	HUTU_KLEAE	P12381	klebsiella
301	5	55.6	8	2	Q9DSN5_9VIRU	Q9dsn5	beet soil-b	374	5	55.6	9	1	IPYR_RHOVI	P82992	rhodopseudo
302	5	55.6	8	2	Q9DSN6_9VIRU	Q9dsn6	beet soil-b	375	5	55.6	9	1	KNL3_BOMVA	P83058	bombina var
303	5	55.6	8	2	Q9E8P7_9VIRU	Q9e8p7	beet soil-b	376	5	55.6	9	1	KNL3_CYPDO	P83659	cyphononyx
304	5	55.6	8	2	Q9E8P8_9VIRU	Q9e8p8	beet soil-b	377	5	55.6	9	1	LMT3_LOCMI	P41489	locusta mig
305	5	55.6	8	2	Q9E8P9_9VIRU	Q9e8p9	beet soil-b	378	5	55.6	9	1	NEUU_CAVPO	P34966	cavia porce
306	5	55.6	8	2	Q9E8Q0_9VIRU	Q9e8q0	beet soil-b	379	5	55.6	9	1	NSK1_SARBU	P41492	sarcophaga
307	5	55.6	8	2	Q9E8Q1_9VIRU	Q9e8q1	beet soil-b	380	5	55.6	9	1	OXYT_CYPCA	P69128	cyprinus ca
308	5	55.6	8	2	Q9E8Q2_9VIRU	Q9e8q2	beet soil-b	381	5	55.6	9	1	OXYT_EISFO	P42998	eisenia foe
309	5	55.6	8	2	Q9E8Q3_9VIRU	Q9e8q3	beet soil-b	382	5	55.6	9	1	OXYT_OCTVU	P80027	octopus vul
310	5	55.6	8	2	Q9E8Q4_9VIRU	Q9e8q4	beet soil-b	383	5	55.6	9	1	OXYT_PETMA	P69129	petromyzon
311	5	55.6	8	2	Q9E8Q5_9VIRU	Q9e8q5	beet soil-b	384	5	55.6	9	1	PGLR_DIAAB	P81179	diaprepes a
312	5	55.6	8	2	Q9E8Q6_9VIRU	Q9e8q6	beet soil-b	385	5	55.6	9	1	PPH1_LYCES	P83380	lycopersico
313	5	55.6	8	2	Q9E8Q7_9VIRU	Q9e8q7	beet soil-b	386	5	55.6	9	1	PPK1_PERAM	P82691	periplaneta
314	5	55.6	8	2	Q9J205_9HEPC	Q9j205	hepatitis c	387	5	55.6	9	1	PVK2_MUSDO	P84355	musca domes
315	5	55.6	8	2	Q56T42_9GEMI	Q56t42	okra yellow	388	5	55.6	9	1	PVK2_SARBU	P84353	sarcophaga
316	5	55.6	8	2	P79940_XENLA	P79940	xenopus lae	389	5	55.6	9	1	PYF2_PENMO	P84006	penaeus mon
317	5	55.6	8	2	Q5YDB9_9PERC	Q5ydb9	xiphister m	390	5	55.6	9	1	PYF3_PENMO	P84007	penaeus mon
318	5	55.6	8	2	Q5YDW3_9PERC	Q5ydw3	xiphister m	391	5	55.6	9	1	PYF4_PENMO	P84008	penaeus mon
319	5	55.6	8	2	Q64IX4_FUNHE	Q64ix4	fundulus he	392	5	55.6	9	1	PYF_LOLVU	P84004	loligo vulg
320	5	55.6	8	2	Q7I5L5_VARDU	Q7i5l5	varanus dum	393	5	55.6	9	1	RS10_SERMA	O68936	serratia ma
321	5	55.6	8	2	Q7LZ46_CTEID	Q7lz46	ctenopharyn	394	5	55.6	9	1	TKC1_CALVO	P41517	calliphora
322	5	55.6	8	2	Q8JFN8_CHICK	Q8jfn8	gallus gall	395	5	55.6	9	1	TKL1_LOCMI	P16223	locusta mig
323	5	55.6	8	2	Q94V82_9SAUR	Q94v82	varanus yuw	396	5	55.6	9	1	TRP4_LEUMA	P81736	leucophaea

397	5	55.6	9	1	ULAK_MOUSE	P99031	mus musculu
398	5	55.6	9	1	UN19_CLOPA	P81355	clostridium
399	5	55.6	9	1	XYLA_STRS8	P19149	streptomyce
400	5	55.6	9	1	YBFR_AZOVI	P25825	azotobacter
401	5	55.6	9	2	Q50832_METVO	Q50832	methanococc
402	5	55.6	9	2	Q7RY89_NEUCR	Q7ry89	neurospora
403	5	55.6	9	2	Q7S182_NEUCR	Q7s182	neurospora
404	5	55.6	9	2	Q7SCD2_NEUCR	Q7scd2	neurospora
405	5	55.6	9	2	Q9UR18_9ASCO	Q9ur18	sclerotium
406	5	55.6	9	2	Q523Q3_MAGGR	Q523q3	magnaporthe
407	5	55.6	9	2	Q14277_HUMAN	Q14277	homo sapien
408	5	55.6	9	2	Q15892_HUMAN	Q15892	homo sapien
409	5	55.6	9	2	Q16220_HUMAN	Q16220	homo sapien
410	5	55.6	9	2	Q16276_HUMAN	Q16276	homo sapien
411	5	55.6	9	2	Q16386_HUMAN	Q16386	homo sapien
412	5	55.6	9	2	Q67AQ6_HUMAN	Q67aq6	homo sapien
413	5	55.6	9	2	Q67AQ7_HUMAN	Q67aq7	homo sapien
414	5	55.6	9	2	Q67AQ8_HUMAN	Q67aq8	homo sapien
415	5	55.6	9	2	Q67AR0_HUMAN	Q67ar0	homo sapien
416	5	55.6	9	2	Q67AR1_HUMAN	Q67ar1	homo sapien
417	5	55.6	9	2	Q67AR4_HUMAN	Q67ar4	homo sapien
418	5	55.6	9	2	Q67AR5_HUMAN	Q67ar5	homo sapien
419	5	55.6	9	2	Q67AR6_HUMAN	Q67ar6	homo sapien
420	5	55.6	9	2	Q67AR7_HUMAN	Q67ar7	homo sapien
421	5	55.6	9	2	Q67AS0_HUMAN	Q67as0	homo sapien
422	5	55.6	9	2	Q67AS3_HUMAN	Q67as3	homo sapien
423	5	55.6	9	2	Q67AT1_HUMAN	Q67at1	homo sapien
424	5	55.6	9	2	Q67AT2_HUMAN	Q67at2	homo sapien
425	5	55.6	9	2	Q67AT6_HUMAN	Q67at6	homo sapien
426	5	55.6	9	2	Q6KER0_HUMAN	Q6ker0	homo sapien
427	5	55.6	9	2	Q6LCV2_HUMAN	Q6lcv2	homo sapien
428	5	55.6	9	2	Q6LE76_HUMAN	Q6le76	homo sapien
429	5	55.6	9	2	Q6LEH2_HUMAN	Q6leh2	homo sapien
430	5	55.6	9	2	Q71UR3_HUMAN	Q71ur3	homo sapien
431	5	55.6	9	2	Q7KYP6_HUMAN	Q7kyp6	homo sapien
432	5	55.6	9	2	Q7M4S2_HUMAN	Q7m4s2	homo sapien
433	5	55.6	9	2	Q7Z4P0_HUMAN	Q7z4p0	homo sapien
434	5	55.6	9	2	Q9BYF9_HUMAN	Q9byf9	homo sapien
435	5	55.6	9	2	Q9H4M8_HUMAN	Q9h4m8	homo sapien
436	5	55.6	9	2	Q9UC36_HUMAN	Q9uc36	homo sapien
437	5	55.6	9	2	Q9UE26_HUMAN	Q9ue26	homo sapien
438	5	55.6	9	2	Q9UKJ6_HUMAN	Q9ukj6	homo sapien
439	5	55.6	9	2	Q9UMA0_HUMAN	Q9uma0	homo sapien
440	5	55.6	9	2	Q9UQA3_HUMAN	Q9uqa3	homo sapien
441	5	55.6	9	2	Q14715_HUMAN	Q14715	homo sapien
442	5	55.6	9	2	Q15999_HUMAN	Q15999	homo sapien
443	5	55.6	9	2	Q27396_BABBO	Q27396	babesia bov
444	5	55.6	9	2	Q5C1F8_SCHJA	Q5c1f8	schistosoma
445	5	55.6	9	2	Q7JNB6_DROVI	Q7jnb6	drosophila
446	5	55.6	9	2	Q7M3L3_PENVA	Q7m3l3	penaeus van
447	5	55.6	9	2	Q7M3N6_GRYBI	Q7m3n6	gryllus bim
448	5	55.6	9	2	Q7M3N7_GRYBI	Q7m3n7	gryllus bim
449	5	55.6	9	2	Q7M3N8_GRYBI	Q7m3n8	gryllus bim
450	5	55.6	9	2	Q9TWD6_LEPDE	Q9twd6	leptinotars
451	5	55.6	9	2	Q9TWV0_ANTEL	Q9twv0	anthopleura
452	5	55.6	9	2	Q9TWX7_MANSE	Q9twx7	manduca sex
453	5	55.6	9	2	P84502_9ANNE	P84502	annelida. l
454	5	55.6	9	2	Q7RE58_PLAYO	Q7re58	plasmodium
455	5	55.6	9	2	Q7RGW1_PLAYO	Q7rgw1	plasmodium
456	5	55.6	9	2	Q8WGE6_PROCL	Q8wge6	procambarus
457	5	55.6	9	2	Q4X9I2_PLACH	Q4x9i2	plasmodium
458	5	55.6	9	2	Q4Y4I3_PLACH	Q4y4i3	plasmodium
459	5	55.6	9	2	Q4Y954_PLACH	Q4y954	plasmodium
460	5	55.6	9	2	Q4Y9N3_PLABE	Q4y9n3	plasmodium
461	5	55.6	9	2	Q6LDS7_RABIT	Q6lds7	oryctolagus
462	5	55.6	9	2	Q7JIS3_LAGOB	Q7jis3	lagenorhync
463	5	55.6	9	2	Q7JIS4_LAGOL	Q7jis4	lagenorhync
464	5	55.6	9	2	Q7JIS5_LAGOL	Q7jis5	lagenorhync
465	5	55.6	9	2	Q7JIS6_LAGOL	Q7jis6	lagenorhync
466	5	55.6	9	2	Q7JIS7_LAGOL	Q7jis7	lagenorhync
467	5	55.6	9	2	Q7JIS8_LAGOL	Q7jis8	lagenorhync
468	5	55.6	9	2	Q7JIS9_LAGAC	Q7jis9	lagenorhync
469	5	55.6	9	2	Q7JIT0_LAGAC	Q7jit0	lagenorhync
470	5	55.6	9	2	Q7JIT1_LAGAC	Q7jit1	lagenorhync
471	5	55.6	9	2	Q7M375_BOVIN	Q7m375	bos taurus
472	5	55.6	9	2	Q8MJN1_CEBPY	Q8mjn1	cebuella py
473	5	55.6	9	2	Q8MJN2_CALJA	Q8mjn2	callithrix
474	5	55.6	9	2	Q8MJN3_CALGO	Q8mjn3	callimico g
475	5	55.6	9	2	Q8MJN4_LEORO	Q8mjn4	leontopithe
476	5	55.6	9	2	Q8MJN5_SAGFU	Q8mjn5	saguinus fu
477	5	55.6	9	2	Q8MJN6_AOTAZ	Q8mjn6	aotus azara
478	5	55.6	9	2	Q8MJN7_SAISC	Q8mjn7	saimiri sci
479	5	55.6	9	2	Q8MJN8_CEBAP	Q8mjn8	cebus apell
480	5	55.6	9	2	Q8MJN9_ATEFU	Q8mjn9	ateles fusc
481	5	55.6	9	2	Q9GJV1_LAGAC	Q9gjv1	lagenorhync
482	5	55.6	9	2	Q9GJV2_LAGOL	Q9gjv2	lagenorhync
483	5	55.6	9	2	Q9GJV3_LAGOB	Q9gjv3	lagenorhync
484	5	55.6	9	2	Q9TRW2_RABIT	Q9trw2	oryctolagus
485	5	55.6	9	2	Q9TT77_BOVIN	Q9tt77	bos taurus
486	5	55.6	9	2	Q9TUY0_MONDO	Q9tuy0	monodelphis
487	5	55.6	9	2	Q9XT05_MACRG	Q9xt05	macropus ru
488	5	55.6	9	2	Q6JDL5_CANFA	Q6jdl5	canis famil
489	5	55.6	9	2	Q4TZV5_PAPHA	Q4tzv5	papio hamad
490	5	55.6	9	2	Q38366_BPPHX	Q38366	bacteriopha
491	5	55.6	9	2	Q42452_WHEAT	Q42452	triticum ae
492	5	55.6	9	2	Q5D4X3_9MYRT	Q5d4x3	lourtella r
493	5	55.6	9	2	Q5D4X7_9MYRT	Q5d4x7	lythrum hys
494	5	55.6	9	2	Q5D4Y6_9MYRT	Q5d4y6	capuronia m
495	5	55.6	9	2	Q5IA44_9ARAE	Q5ia44	philodendro
496	5	55.6	9	2	Q5IA45_MAHBE	Q5ia45	mahonia bea
497	5	55.6	9	2	Q5VB50_9FILI	Q5vb50	bolbitis au
498	5	55.6	9	2	Q5VB51_9FILI	Q5vb51	elaphogloss
499	5	55.6	9	2	Q5VB52_9FILI	Q5vb52	elaphogloss
500	5	55.6	9	2	Q5VB53_9FILI	Q5vb53	elaphogloss
501	5	55.6	9	2	Q5VB54_9FILI	Q5vb54	elaphogloss
502	5	55.6	9	2	Q5VB55_9FILI	Q5vb55	elaphogloss
503	5	55.6	9	2	Q5VB56_9FILI	Q5vb56	elaphogloss
504	5	55.6	9	2	Q5VB57_9FILI	Q5vb57	elaphogloss
505	5	55.6	9	2	Q5VB58_9FILI	Q5vb58	elaphogloss
506	5	55.6	9	2	Q5VB59_9FILI	Q5vb59	elaphogloss
507	5	55.6	9	2	Q5VB60_9FILI	Q5vb60	elaphogloss
508	5	55.6	9	2	Q5VB61_9FILI	Q5vb61	elaphogloss
509	5	55.6	9	2	Q5VB62_9FILI	Q5vb62	elaphogloss
510	5	55.6	9	2	Q5VB63_9FILI	Q5vb63	elaphogloss
511	5	55.6	9	2	Q5VB64_9FILI	Q5vb64	elaphogloss
512	5	55.6	9	2	Q5VB65_9FILI	Q5vb65	elaphogloss
513	5	55.6	9	2	Q5VB66_9FILI	Q5vb66	elaphogloss
514	5	55.6	9	2	Q5VB67_9FILI	Q5vb67	elaphogloss
515	5	55.6	9	2	Q5VB68_9FILI	Q5vb68	elaphogloss
516	5	55.6	9	2	Q5VB69_9FILI	Q5vb69	elaphogloss
517	5	55.6	9	2	Q5VB70_9FILI	Q5vb70	elaphogloss
518	5	55.6	9	2	Q5VB71_9FILI	Q5vb71	elaphogloss
519	5	55.6	9	2	Q5VB72_9FILI	Q5vb72	elaphogloss
520	5	55.6	9	2	Q5VB73_9FILI	Q5vb73	elaphogloss
521	5	55.6	9	2	Q5VB74_9FILI	Q5vb74	elaphogloss
522	5	55.6	9	2	Q5VB75_9FILI	Q5vb75	elaphogloss
523	5	55.6	9	2	Q5VB76_9FILI	Q5vb76	elaphogloss
524	5	55.6	9	2	Q5VB77_9FILI	Q5vb77	elaphogloss
525	5	55.6	9	2	Q5VB78_9FILI	Q5vb78	elaphogloss
526	5	55.6	9	2	Q5VB79_9FILI	Q5vb79	elaphogloss
527	5	55.6	9	2	Q5VB80_9FILI	Q5vb80	elaphogloss
528	5	55.6	9	2	Q5VB81_9FILI	Q5vb81	elaphogloss
529	5	55.6	9	2	Q5VB82_9FILI	Q5vb82	elaphogloss
530	5	55.6	9	2	Q5VB83_9FILI	Q5vb83	elaphogloss
531	5	55.6	9	2	Q5VB84_9FILI	Q5vb84	elaphogloss
532	5	55.6	9	2	Q5VB85_9FILI	Q5vb85	elaphogloss
533	5	55.6	9	2	Q5VB86_9FILI	Q5vb86	elaphogloss
534	5	55.6	9	2	Q5VB87_9FILI	Q5vb87	elaphogloss
535	5	55.6	9	2	Q5VB88_9FILI	Q5vb88	elaphogloss
536	5	55.6	9	2	Q5VB89_9FILI	Q5vb89	elaphogloss
537	5	55.6	9	2	Q5VB90_9FILI	Q5vb90	elaphogloss
538	5	55.6	9	2	Q5VB91_9FILI	Q5vb91	elaphogloss
539	5	55.6	9	2	Q5VB92_9FILI	Q5vb92	elaphogloss
540	5	55.6	9	2	Q5VB93_9FILI	Q5vb93	elaphogloss
541	5	55.6	9	2	Q5VB94_9FILI	Q5vb94	elaphogloss
542	5	55.6	9	2	Q5VB95_9FILI	Q5vb95	elaphogloss

543	5	55.6	9	2	Q5VB96_9FILI	Q5vb96	elaphogloss
544	5	55.6	9	2	Q5VB97_9FILI	Q5vb97	elaphogloss
545	5	55.6	9	2	Q5VB98_9FILI	Q5vb98	elaphogloss
546	5	55.6	9	2	Q5VB99_9FILI	Q5vb99	elaphogloss
547	5	55.6	9	2	Q5VBA0_9FILI	Q5vba0	elaphogloss
548	5	55.6	9	2	Q5VBA1_9FILI	Q5vba1	elaphogloss
549	5	55.6	9	2	Q5VBA2_9FILI	Q5vba2	elaphogloss
550	5	55.6	9	2	Q5VBA3_9FILI	Q5vba3	elaphogloss
551	5	55.6	9	2	Q5VBA4_9FILI	Q5vba4	elaphogloss
552	5	55.6	9	2	Q5VBA5_9FILI	Q5vba5	elaphogloss
553	5	55.6	9	2	Q5VBA6_9FILI	Q5vba6	elaphogloss
554	5	55.6	9	2	Q5VBA7_9FILI	Q5vba7	elaphogloss
555	5	55.6	9	2	Q5VBA8_9FILI	Q5vba8	elaphogloss
556	5	55.6	9	2	Q5VBA9_9FILI	Q5vba9	elaphogloss
557	5	55.6	9	2	Q5VBB0_9FILI	Q5vbb0	elaphogloss
558	5	55.6	9	2	Q5VBB1_9FILI	Q5vbb1	elaphogloss
559	5	55.6	9	2	Q5VBB2_9FILI	Q5vbb2	elaphogloss
560	5	55.6	9	2	Q5VBB3_9FILI	Q5vbb3	elaphogloss
561	5	55.6	9	2	Q5VBB4_9FILI	Q5vbb4	elaphogloss
562	5	55.6	9	2	Q5VBB5_9FILI	Q5vbb5	elaphogloss
563	5	55.6	9	2	Q5VBB6_9FILI	Q5vbb6	elaphogloss
564	5	55.6	9	2	Q5VBB7_9FILI	Q5vbb7	elaphogloss
565	5	55.6	9	2	Q5VBB8_9FILI	Q5vbb8	elaphogloss
566	5	55.6	9	2	Q5VBB9_9FILI	Q5vbb9	elaphogloss
567	5	55.6	9	2	Q5VBC0_9FILI	Q5vbc0	elaphogloss
568	5	55.6	9	2	Q5VBC1_9FILI	Q5vbc1	elaphogloss
569	5	55.6	9	2	Q5VBC2_9FILI	Q5vbc2	elaphogloss
570	5	55.6	9	2	Q5VBC3_9FILI	Q5vbc3	elaphogloss
571	5	55.6	9	2	Q5VBC4_9FILI	Q5vbc4	elaphogloss
572	5	55.6	9	2	Q5VBC5_9FILI	Q5vbc5	elaphogloss
573	5	55.6	9	2	Q5VBC6_9FILI	Q5vbc6	elaphogloss
574	5	55.6	9	2	Q5VBC7_9FILI	Q5vbc7	elaphogloss
575	5	55.6	9	2	Q5VBC8_9FILI	Q5vbc8	elaphogloss
576	5	55.6	9	2	Q5VBC9_9FILI	Q5vbc9	elaphogloss
577	5	55.6	9	2	Q5Y9F5_9BRYO	Q5y9f5	callier gon
578	5	55.6	9	2	Q672I8_PENAM	Q672i8	pennisetum
579	5	55.6	9	2	Q6A1H7_HORVD	Q6a1h7	hordeum vul
580	5	55.6	9	2	Q6EUV8_GERHY	Q6euv8	gerbera hyb
581	5	55.6	9	2	Q6EX64_9LAMI	Q6ex64	hyptis flor
582	5	55.6	9	2	Q6RVW6_CAPAN	Q6rvw6	capsicum an
583	5	55.6	9	2	Q6VR25_THECC	Q6vr25	theobroma c
584	5	55.6	9	2	Q6XBN2_9BRYO	Q6xbn2	vittia pach
585	5	55.6	9	2	Q6ZZ00_9CARY	Q6zz00	silene rotu
586	5	55.6	9	2	Q70Y63_CONTO	Q70y63	congea tome
587	5	55.6	9	2	Q70Y76_SOLSC	Q70y76	solenostemo
588	5	55.6	9	2	Q70Y83_9LAMI	Q70y83	plectranthu
589	5	55.6	9	2	Q7EXP6_HORVD	Q7exp6	hordeum vul
590	5	55.6	9	2	Q7X8P7_MAIZE	Q7x8p7	zea mays (m
591	5	55.6	9	2	Q85G96_9BRYO	Q85g96	pyrrhobryum
592	5	55.6	9	2	Q85V64_EUCGR	Q85v64	eucalyptus
593	5	55.6	9	2	Q8MBF4_IPOQU	Q8mbf4	ipomoea qua
594	5	55.6	9	2	Q8MEM3_9ROSI	Q8mem3	howittia tr
595	5	55.6	9	2	Q940K4_ARATH	Q940k4	arabidopsis
596	5	55.6	9	2	Q95GN1_PELHO	Q95gn1	pelargonium
597	5	55.6	9	2	Q9AXH8_MESCR	Q9axh8	mesembryant
598	5	55.6	9	2	Q9FEC0_HORVU	Q9fec0	hordeum vul
599	5	55.6	9	2	Q9GCV6_9LILI	Q9gcv6	sclerosperm
600	5	55.6	9	2	Q9GD12_9LILI	Q9gd12	linospadix
601	5	55.6	9	2	Q9S8J8_ORYSA	Q9s8j8	oryza sativ
602	5	55.6	9	2	Q9TKD9_9MYRT	Q9tkd9	pericalymma
603	5	55.6	9	2	Q9TKF2_9MYRT	Q9tkf2	asteromyrtu
604	5	55.6	9	2	Q9TKG1_9MYRT	Q9tkg1	calothamnus
605	5	55.6	9	2	Q9TLM7_9FLOR	Q9tlm7	laurencia v
606	5	55.6	9	2	Q5SBX0_9POAL	Q5sbx0	restio insi
607	5	55.6	9	2	Q4QWV3_9MARC	Q4qwv3	pallavicini
608	5	55.6	9	2	Q4QWV9_9MARC	Q4qwv9	jensenia sp
609	5	55.6	9	2	Q4U0F2_MALDO	Q4u0f2	malus domes
610	5	55.6	9	2	O30790_ERWAM	O30790	erwinia amy
611	5	55.6	9	2	P72345_PSESX	P72345	pseudomonas
612	5	55.6	9	2	Q45852_CLOBU	Q45852	clostridium
613	5	55.6	9	2	Q47410_ECOLI	Q47410	escherichia
614	5	55.6	9	2	Q53914_9ACTO	Q53914	streptomyce
615	5	55.6	9	2	Q5K2V4_9NOST	Q5k2v4	nodularia h
616	5	55.6	9	2	Q5K2V7_NODSP	Q5k2v7	nodularia s
617	5	55.6	9	2	P82568_STRPY	P82568	streptococc
618	5	55.6	9	2	P83539_LACSN	P83539	lactobacill
619	5	55.6	9	2	Q6LDL7_STAAU	Q6ldl7	staphylococ
620	5	55.6	9	2	Q6VCX0_9ACTO	Q6vcx0	streptomyce
621	5	55.6	9	2	Q6VFQ2_VIBFI	Q6vfq2	vibrio fisc
622	5	55.6	9	2	Q711U9_9RHIZ	Q711u9	azorhizobiu
623	5	55.6	9	2	Q711V3_AZOCA	Q711v3	azorhizobiu
624	5	55.6	9	2	Q712B7_SINTE	Q712b7	sinorhizobi
625	5	55.6	9	2	Q71UF6_YEREN	Q71uf6	yersinia en
626	5	55.6	9	2	Q798K5_STRLI	Q798k5	streptomyce
627	5	55.6	9	2	Q7M0L7_9STRE	Q7m0l7	streptococc
628	5	55.6	9	2	Q7M151_9BACT	Q7m151	unidentifie
629	5	55.6	9	2	Q8GIZ6_LACDL	Q8giz6	lactobacill
630	5	55.6	9	2	Q8RKU3_BORBU	Q8rku3	borrelia bu
631	5	55.6	9	2	Q93EZ0_STRAG	Q93ez0	streptococc
632	5	55.6	9	2	Q9JNI6_STRPY	Q9jni6	streptococc
633	5	55.6	9	2	Q9R7H9_HAEIN	Q9r7h9	haemophilus
634	5	55.6	9	2	Q51765_PSEFL	Q51765	pseudomonas
635	5	55.6	9	2	O08979_MOUSE	O08979	mus musculu
636	5	55.6	9	2	Q6LAQ1_RAT	Q6laq1	rattus norv
637	5	55.6	9	2	Q6YF34_RAT	Q6yf34	rattus norv
638	5	55.6	9	2	Q78E72_RAT	Q78e72	rattus norv
639	5	55.6	9	2	Q7M078_RAT	Q7m078	rattus norv
640	5	55.6	9	2	Q8R514_RAT	Q8r514	rattus norv
641	5	55.6	9	2	Q920I2_MOUSE	Q920i2	mus musculu
642	5	55.6	9	2	Q9QVH9_9MURI	Q9qvh9	mus sp. sup
643	5	55.6	9	2	Q9QZA7_MOUSE	Q9qza7	mus musculu
644	5	55.6	9	2	Q9QWG2_MOUSE	Q9qwg2	mus musculu
645	5	55.6	9	2	O88889_MOUSE	O88889	mus musculu
646	5	55.6	9	2	Q61723_MOUSE	Q61723	mus musculu
647	5	55.6	9	2	Q62530_MUSSP	Q62530	mus spretus
648	5	55.6	9	2	Q4W8Q9_MOUSE	Q4w8q9	mus musculu
649	5	55.6	9	2	O71066_9PARA	O71066	canine dist
650	5	55.6	9	2	O71067_9PARA	O71067	canine dist
651	5	55.6	9	2	O71068_9PARA	O71068	canine dist
652	5	55.6	9	2	P90359_9POTY	P90359	barley mild
653	5	55.6	9	2	Q66545_9GAMA	Q66545	human herpe
654	5	55.6	9	2	Q67605_SLCV	Q67605	squash leaf
655	5	55.6	9	2	Q67606_SLCV	Q67606	squash leaf
656	5	55.6	9	2	Q69349_HHV2	Q69349	human herpe
657	5	55.6	9	2	Q69473_HHV1	Q69473	human herpe
658	5	55.6	9	2	Q6T1E2_CVHSA	Q6t1e2	sars corona
659	5	55.6	9	2	Q82622_9CORO	Q82622	avian infec
660	5	55.6	9	2	Q83622_9FLAV	Q83622	murray vall
661	5	55.6	9	2	Q84333_SV40	Q84333	simian viru
662	5	55.6	9	2	Q88953_9POXV	Q88953	vaccinia vi
663	5	55.6	9	2	Q8QVD3_9MONO	Q8qvd3	ovine respi
664	5	55.6	9	2	Q9E1U7_HPBV0	Q9e1u7	hepatitis b
665	5	55.6	9	2	Q9IBM8_9POLY	Q9ibm8	simian viru
666	5	55.6	9	2	Q9PYK1_9POLY	Q9pyk1	simian viru
667	5	55.6	9	2	P84497_TRASC	P84497	trachemys s
668	5	55.6	9	2	Q5YDV1_9PERC	Q5ydv1	xiphister m
669	5	55.6	9	2	Q64IX3_FUNHE	Q64ix3	fundulus he
670	5	55.6	9	2	Q673X5_9CORV	Q673x5	platysteira
671	5	55.6	9	2	Q673Y1_9CORV	Q673y1	lanioturdus
672	5	55.6	9	2	Q673Y6_9CORV	Q673y6	platysteira
673	5	55.6	9	2	Q673Z9_9CORV	Q673z9	batis poens
674	5	55.6	9	2	Q691D6_ANOSA	Q691d6	anolis sagr
675	5	55.6	9	2	Q6HA69_TRITG	Q6ha69	trimeresuru
676	5	55.6	9	2	Q6HA76_9SAUR	Q6ha76	trimeresuru
677	5	55.6	9	2	Q71DX2_9SAUR	Q71dx2	urostrophus
678	5	55.6	9	2	Q75QC9_GALLA	Q75qc9	gallus lafa
679	5	55.6	9	2	Q7LZ50_CHICK	Q7lz50	gallus gall
680	5	55.6	9	2	Q7LZ66_MELGA	Q7lz66	meleagris g
681	5	55.6	9	2	Q7LZJ8_RANTE	Q7lzj8	rana tempor
682	5	55.6	9	2	Q8AYL5_CARAU	Q8ayl5	carassius a
683	5	55.6	9	2	Q8SHF0_CHANA	Q8shf0	chamaeleo n
684	5	55.6	9	2	Q92009_CHICK	Q92009	gallus gall
685	5	55.6	9	2	Q94VC6_9SAUR	Q94vc6	varanus pil
686	5	55.6	9	2	Q94VD8_VARNI	Q94vd8	varanus nil
687	5	55.6	9	2	Q94VE1_VARME	Q94ve1	varanus mer
688	5	55.6	9	2	Q94VG2_VARIN	Q94vg2	varanus ind

689	5	55.6	9	2	Q94VH4_9SAUR	Q94vh4 varanus gla
690	5	55.6	9	2	Q94VI0_VARGI	Q94vi0 varanus gig
691	5	55.6	9	2	Q94VI8_VARER	Q94vi8 varanus ere
692	5	55.6	9	2	Q94VJ1_VARDO	Q94vj1 varanus dor
693	5	55.6	9	2	Q9IAI4_CHICK	Q9iai4 gallus gall
694	5	55.6	9	2	Q9PRJ4_LEPOS	Q9prj4 lepisosteus
695	5	55.6	9	2	Q9T688_GECGE	Q9t688 gecko gecko
696	5	55.6	9	2	Q50GN1_CHICK	Q50gn1 gallus gall
697	5	55.6	9	2	Q4PU39_9CICH	Q4pu39 lepidiolamp
698	5	55.6	9	2	Q7LZI7_9NEOB	Q7lzi7 heleophryne
699	5	55.6	9	2	O12096_CAEV	O12096 caprine art
700	5	55.6	9	2	O12098_CAEV	O12098 caprine art
701	5	55.6	9	2	O12100_CAEV	O12100 caprine art
702	5	55.6	9	2	O12102_CAEV	O12102 caprine art
703	5	55.6	9	2	O12104_CAEV	O12104 caprine art
704	5	55.6	9	2	Q64972_AVEVR	Q64972 avian rous-
705	5	55.6	9	2	Q70140_9HIV1	Q70140 human immun
706	5	55.6	9	2	Q85599_MLVMO	Q85599 moloney mur
707	5	55.6	9	2	Q8UTD7_9HIV1	Q8utd7 human immun
708	5	55.6	9	2	Q8AEW8_9HIV1	Q8aew8 human immun
709	5	55.6	10	1	AKHX_LOCMI	P81626 locusta mig
710	5	55.6	10	1	ANGT1_BOTJA	Q10581 bothrops ja
711	5	55.6	10	1	ANGT_BOVIN	P01017 bos taurus
712	5	55.6	10	1	ANGT_CHICK	P67885 gallus gall
713	5	55.6	10	1	ANGT_COTJA	P67886 coturnix co
714	5	55.6	10	1	ANOP_ANOSM	P0c005 anoplius sa
715	5	55.6	10	1	APE_CAPGI	P80474 capnocytoph
716	5	55.6	10	1	BPP2_BOTJA	P01022 bothrops ja
717	5	55.6	10	1	BRK_ONCMY	Q9prz1 oncorhynchu
718	5	55.6	10	1	CATB_SHEEP	P83205 ovis aries
719	5	55.6	10	1	COX81_CANFA	P61904 canis famil
720	5	55.6	10	1	COX81_RABIT	P80336 oryctolagus
721	5	55.6	10	1	COX82_CANFA	P61905 canis famil
722	5	55.6	10	1	FAR2_PENMO	P83317 penaeus mon
723	5	55.6	10	1	FAR5_MACRS	P83278 macrobrachi
724	5	55.6	10	1	FAR6_PANRE	P82660 panagrellus
725	5	55.6	10	1	FAR7_MACRS	P83280 macrobrachi
726	5	55.6	10	1	FARC_CALVO	P41867 calliphora
727	5	55.6	10	1	FARP_LOCMI	P84306 locusta mig
728	5	55.6	10	1	FARP_MANSE	P18523 manduca sex
729	5	55.6	10	1	FARP_MYTED	P42560 mytilus edu
730	5	55.6	10	1	FARP_SCHGR	P84307 schistocerc
731	5	55.6	10	1	FIBB_CERSI	P14537 ceratotheri
732	5	55.6	10	1	GS09_BACSU	P80243 bacillus su
733	5	55.6	10	1	LCMS_LEUMA	P21144 leucophaea
734	5	55.6	10	1	LPK2_LOCMI	P41488 locusta mig
735	5	55.6	10	1	LSK2_LEUMA	P67802 leucophaea
736	5	55.6	10	1	LSK2_PERAM	P67803 periplaneta
737	5	55.6	10	1	MALE_KLEPN	Q05564 klebsiella
738	5	55.6	10	1	MP2_MICOC	P81533 microplitis
739	5	55.6	10	1	NEMS_SARBU	P61850 sarcophaga
740	5	55.6	10	1	PA66D_SHEEP	P83496 ovis aries
741	5	55.6	10	1	PORB_METTM	P80901 methanobact
742	5	55.6	10	1	PSBF_CAPAN	Q03367 capsicum an
743	5	55.6	10	1	PVK_LOCMI	P83382 locusta mig
744	5	55.6	10	1	PVK_PHYMO	P84442 phymatéus m
745	5	55.6	10	1	Q2OB_COMTE	P80465 comamonas t
746	5	55.6	10	1	RL16_ACHLA	P29221 acholeplasm
747	5	55.6	10	1	RNAM_PLESA	P84528 pleurotus s
748	5	55.6	10	1	RT02_BOVIN	P82923 bos taurus
749	5	55.6	10	1	TKL2_LOCMI	P16224 locusta mig
750	5	55.6	10	1	TKL3_LOCMI	P30249 locusta mig
751	5	55.6	10	1	TKL4_LOCMI	P30250 locusta mig
752	5	55.6	10	1	TKN_PHYBI	P08610 phyllomedus
753	5	55.6	10	1	TKU1_UREUN	P40751 urechis uni
754	5	55.6	10	1	TKU2_UREUN	P40752 urechis uni
755	5	55.6	10	1	TPIS_NICPL	P19118 nicotiana p
756	5	55.6	10	1	TRP5_LEUMA	P81737 leucophaea
757	5	55.6	10	1	TRP6_LEUMA	P81738 leucophaea
758	5	55.6	10	1	TRP7_LEUMA	P81739 leucophaea
759	5	55.6	10	1	TRP8_LEUMA	P81740 leucophaea
760	5	55.6	10	1	TRP9_LEUMA	P81741 leucophaea
761	5	55.6	10	1	UPA5_HUMAN	P30091 homo sapien
762	5	55.6	10	1	UPA9_HUMAN	P30095 homo sapien
763	5	55.6	10	2	Q7M530_PYRFU	Q7m530 pyrococcus
764	5	55.6	10	2	Q7M4X7_FUSSP	Q7m4x7 fusarium sp
765	5	55.6	10	2	Q7RYS0_NEUCR	Q7rys0 neurospora
766	5	55.6	10	2	Q7RZZ2_NEUCR	Q7rzz2 neurospora
767	5	55.6	10	2	Q7S184_NEUCR	Q7s184 neurospora
768	5	55.6	10	2	Q7S377_NEUCR	Q7s377 neurospora
769	5	55.6	10	2	Q7S5J5_NEUCR	Q7s5j5 neurospora
770	5	55.6	10	2	Q7SA62_NEUCR	Q7sa62 neurospora
771	5	55.6	10	2	Q9C1K4_GLOMO	Q9c1k4 glomus moss
772	5	55.6	10	2	Q15342_HUMAN	Q15342 homo sapien
773	5	55.6	10	2	Q6LB24_HUMAN	Q6lb24 homo sapien
774	5	55.6	10	2	Q6LCI4_HUMAN	Q6lci4 homo sapien
775	5	55.6	10	2	Q6LEM6_HUMAN	Q6lem6 homo sapien
776	5	55.6	10	2	Q6MZE6_HUMAN	Q6mze6 homo sapien
777	5	55.6	10	2	Q70LT3_HUMAN	Q70lt3 homo sapien
778	5	55.6	10	2	Q712L8_HUMAN	Q712l8 homo sapien
779	5	55.6	10	2	Q7KZ15_HUMAN	Q7kz15 homo sapien
780	5	55.6	10	2	Q7Z5A2_HUMAN	Q7z5a2 homo sapien
781	5	55.6	10	2	Q8N6B1_HUMAN	Q8n6b1 homo sapien
782	5	55.6	10	2	Q8NEY9_HUMAN	Q8ney9 homo sapien
783	5	55.6	10	2	Q8WXB5_HUMAN	Q8wxb5 homo sapien
784	5	55.6	10	2	Q9UCQ4_HUMAN	Q9ucq4 homo sapien
785	5	55.6	10	2	Q9UCU6_HUMAN	Q9ucu6 homo sapien
786	5	55.6	10	2	Q9UDE8_HUMAN	Q9ude8 homo sapien
787	5	55.6	10	2	Q9UNF2_HUMAN	Q9unf2 homo sapien
788	5	55.6	10	2	Q53TW5_HUMAN	Q53tw5 homo sapien
789	5	55.6	10	2	Q5S4Q0_HUMAN	Q5s4q0 homo sapien
790	5	55.6	10	2	Q86XP4_HUMAN	Q86xp4 homo sapien
791	5	55.6	10	2	Q4ZFV9_HUMAN	Q4zfv9 homo sapien
792	5	55.6	10	2	Q5JPL1_HUMAN	Q5jpl1 homo sapien
793	5	55.6	10	2	Q35013_MELJA	Q35013 meloidogyne
794	5	55.6	10	2	Q5BRP3_SCHJA	Q5brp3 schistosoma
795	5	55.6	10	2	Q5C1M9_SCHJA	Q5c1m9 schistosoma
796	5	55.6	10	2	Q5C4A5_SCHJA	Q5c4a5 schistosoma
797	5	55.6	10	2	P82383_DROME	P82383 drosophila
798	5	55.6	10	2	Q7M4C2_ECHMA	Q7m4c2 echinometra
799	5	55.6	10	2	Q86D30_TRYCR	Q86d30 trypanosoma
800	5	55.6	10	2	Q8WPL6_9UROC	Q8wpl6 oikopleura
801	5	55.6	10	2	Q9TWU1_FUSFE	Q9twu1 fusinus fer
802	5	55.6	10	2	Q9TWX1_DERFA	Q9twx1 dermatophag
803	5	55.6	10	2	Q7RBG5_PLAYO	Q7rbg5 plasmodium
804	5	55.6	10	2	Q7RDS6_PLAYO	Q7rds6 plasmodium
805	5	55.6	10	2	Q7RKS4_PLAYO	Q7rks4 plasmodium
806	5	55.6	10	2	Q4XE58_PLACH	Q4xe58 plasmodium
807	5	55.6	10	2	Q4XZ43_PLACH	Q4xz43 plasmodium
808	5	55.6	10	2	Q4Z638_PLABE	Q4z638 plasmodium
809	5	55.6	10	2	Q5QSI4_DIDMR	Q5qsi4 didelphis m
810	5	55.6	10	2	Q6IG13_BOVIN	Q6ig13 bos taurus
811	5	55.6	10	2	Q7M2U1_BOVIN	Q7m2u1 bos taurus
812	5	55.6	10	2	Q8SPN8_MACMU	Q8spn8 macaca mula
813	5	55.6	10	2	Q95H99_PAPAN	Q95h99 papio anubi
814	5	55.6	10	2	Q95HF4_PAPAN	Q95hf4 papio anubi
815	5	55.6	10	2	Q95HF5_PAPAN	Q95hf5 papio anubi
816	5	55.6	10	2	Q95HF6_PAPAN	Q95hf6 papio anubi
817	5	55.6	10	2	Q95HF7_PAPAN	Q95hf7 papio anubi
818	5	55.6	10	2	Q95HF8_PAPAN	Q95hf8 papio anubi
819	5	55.6	10	2	Q95HF9_PAPAN	Q95hf9 papio anubi
820	5	55.6	10	2	Q95HG0_PAPAN	Q95hg0 papio anubi
821	5	55.6	10	2	Q95M70_TRIVU	Q95m70 trichosurus
822	5	55.6	10	2	Q9GKI4_MACAR	Q9gki4 macaca arct
823	5	55.6	10	2	Q9GKI5_PANTR	Q9gki5 pan troglod
824	5	55.6	10	2	Q9TQV4_HORSE	Q9tqv4 e equus cab
825	5	55.6	10	2	Q9TRS3_PIG	Q9trs3 sus scrofa
826	5	55.6	10	2	Q9XS84_HORSE	Q9xs84 equus cabal
827	5	55.6	10	2	Q6JDL6_CANFA	Q6jdl6 canis famil
828	5	55.6	10	2	Q7M2N0_BOVIN	Q7m2n0 bos taurus
829	5	55.6	10	2	Q7M3E8_PIG	Q7m3e8 sus scrofa
830	5	55.6	10	2	Q38217_9CAUD	Q38217 lactococcus
831	5	55.6	10	2	Q5D4Y2_9MYRT	Q5d4y2 ginoria ame
832	5	55.6	10	2	Q5I8T2_CAMSI	Q5i8t2 camellia si
833	5	55.6	10	2	Q5I8T5_9ERIC	Q5i8t5 camellia te
834	5	55.6	10	2	Q5I8T6_9ERIC	Q5i8t6 camellia fu

835	5	55.6	10	2	Q5I8T7_CAMSI	Q5i8t7	camellia si	908	5	55.6	10	2	P82588_STRPY	P82588	streptococc
836	5	55.6	10	2	P82434_TOBAC	P82434	nicotiana t	909	5	55.6	10	2	P83066_BACCE	P83066	bacillus ce
837	5	55.6	10	2	P82443_TOBAC	P82443	nicotiana t	910	5	55.6	10	2	P83067_BACCE	P83067	bacillus ce
838	5	55.6	10	2	P82937_HORVU	P82937	hordeum vul	911	5	55.6	10	2	P83160_ANASL	P83160	anabaena sp
839	5	55.6	10	2	P82938_HORVU	P82938	hordeum vul	912	5	55.6	10	2	Q7M0J3_MARPU	Q7m0j3	marichromat
840	5	55.6	10	2	Q5Y9E0_9BRYO	Q5y9e0	scorpidium	913	5	55.6	10	2	Q7M0M6_DESDE	Q7m0m6	desulfovibr
841	5	55.6	10	2	Q5Y9F4_9BRYO	Q5y9f4	calliergon	914	5	55.6	10	2	Q7M0N4_SERMA	Q7m0n4	serratia ma
842	5	55.6	10	2	Q6JVP0_WOLBI	Q6jvp0	wollastonia	915	5	55.6	10	2	Q847B5_BACAM	Q847b5	bacillus am
843	5	55.6	10	2	Q6JVP3_9ASTR	Q6jvp3	otopappus e	916	5	55.6	10	2	Q8KHN9_CLOBO	Q8khn9	clostridium
844	5	55.6	10	2	Q6JVP6_9ASTR	Q6jvp6	elaphandra	917	5	55.6	10	2	Q8RJF1_PSEFL	Q8rjf1	pseudomonas
845	5	55.6	10	2	Q6JVP8_9ASTR	Q6jvp8	dimerostemm	918	5	55.6	10	2	Q93LX4_VIBCH	Q93lx4	vibrio chol
846	5	55.6	10	2	Q6JVQ0_9ASTR	Q6jvq0	dimerostemm	919	5	55.6	10	2	Q93UU2_ECO57	Q93uu2	escherichia
847	5	55.6	10	2	Q6JVQ3_9ASTR	Q6jvq3	baltimora r	920	5	55.6	10	2	Q9JNC9_STRPY	Q9jnc9	streptococc
848	5	55.6	10	2	Q6JVQ5_9ASTR	Q6jvq5	angelphytum	921	5	55.6	10	2	Q9R5N1_CLOBO	Q9r5n1	clostridium
849	5	55.6	10	2	Q6JVQ7_9ASTR	Q6jvq7	angelphytum	922	5	55.6	10	2	Q9R5N3_CLOBO	Q9r5n3	clostridium
850	5	55.6	10	2	Q6KC69_EUCGL	Q6kc69	eucalyptus	923	5	55.6	10	2	Q9R7J9_HELPY	Q9r7j9	helicobacte
851	5	55.6	10	2	Q70Y78_9LAMI	Q70y78	plectranthu	924	5	55.6	10	2	Q9S3J6_ECOLI	Q9s3j6	escherichia
852	5	55.6	10	2	Q7M1F6_9POAL	Q7m1f6	haynaldia v	925	5	55.6	10	2	Q9X533_ECOLI	Q9x533	escherichia
853	5	55.6	10	2	Q7M1F7_9POAL	Q7m1f7	haynaldia v	926	5	55.6	10	2	Q9X534_9ENTR	Q9x534	leclercia a
854	5	55.6	10	2	Q7M1I1_PHAVU	Q7m1i1	phaseolus v	927	5	55.6	10	2	O54217_STAEP	O54217	staphylococ
855	5	55.6	10	2	Q7M1I6_TRIKI	Q7m1i6	trichosanth	928	5	55.6	10	2	Q71V02_PSEAE	Q71v02	pseudomonas
856	5	55.6	10	2	Q7M1V8_NICPL	Q7m1v8	nicotiana p	929	5	55.6	10	2	Q6JL97_NEIGO	Q6jl97	neisseria g
857	5	55.6	10	2	Q7M278_TRITU	Q7m278	triticum tu	930	5	55.6	10	2	Q8VN85_HELPY	Q8vn85	helicobacte
858	5	55.6	10	2	Q7M282_ORYSA	Q7m282	oryza sativ	931	5	55.6	10	2	P74843_STRTH	P74843	streptomyce
859	5	55.6	10	2	Q7M2G1_VICFA	Q7m2g1	vicia faba	932	5	55.6	10	2	Q48469_KLEPN	Q48469	klebsiella
860	5	55.6	10	2	Q7M2G2_VICFA	Q7m2g2	vicia faba	933	5	55.6	10	2	Q79AV7_KLEPN	Q79av7	klebsiella
861	5	55.6	10	2	Q85AZ9_9BRYO	Q85az9	pyrrhobryum	934	5	55.6	10	2	Q5D4R6_9BURK	Q5d4r6	alcaligenes
862	5	55.6	10	2	Q85BV5_EUCGR	Q85bv5	eucalyptus	935	5	55.6	10	2	Q7WUG1_PSEFL	Q7wug1	pseudomonas
863	5	55.6	10	2	Q85BV6_EUCGR	Q85bv6	eucalyptus	936	5	55.6	10	2	O08622_RAT	O08622	rattus norv
864	5	55.6	10	2	Q85BV7_EUCGR	Q85bv7	eucalyptus	937	5	55.6	10	2	Q63389_RAT	Q63389	rattus norv
865	5	55.6	10	2	Q85V65_EUCGR	Q85v65	eucalyptus	938	5	55.6	10	2	Q68SM8_CHAPN	Q68sm8	chaetodipus
866	5	55.6	10	2	Q85V66_EUCGR	Q85v66	eucalyptus	939	5	55.6	10	2	Q68SM9_THOMO	Q68sm9	thomomys mo
867	5	55.6	10	2	Q85V67_EUCGR	Q85v67	eucalyptus	940	5	55.6	10	2	Q68SN0_CRATY	Q68sn0	cratogeomys
868	5	55.6	10	2	Q8GZC8_HORVU	Q8gzc8	hordeum vul	941	5	55.6	10	2	Q68SN1_CRAGY	Q68sn1	cratogeomys
869	5	55.6	10	2	Q8HUB4_9BRYO	Q8hub4	anomobryum	942	5	55.6	10	2	Q6LD58_9MURI	Q6ld58	mus sp. alp
870	5	55.6	10	2	Q8MAZ9_9ASTE	Q8maz9	dicranostyl	943	5	55.6	10	2	Q6S510_MOUSE	Q6s510	mus musculu
871	5	55.6	10	2	Q8MBB7_9ASTE	Q8mbb7	merremia ae	944	5	55.6	10	2	Q80WD4_MOUSE	Q80wd4	mus musculu
872	5	55.6	10	2	Q8SAC2_9BRYO	Q8sac2	amblystegiu	945	5	55.6	10	2	Q80WD9_9MURI	Q80wd9	rattus sp.
873	5	55.6	10	2	Q947R7_SOLTU	Q947r7	solanum tub	946	5	55.6	10	2	Q80Z98_RAT	Q80z98	rattus norv
874	5	55.6	10	2	Q94II9_MAIZE	Q94ii9	zea mays (m	947	5	55.6	10	2	Q8BHN2_MOUSE	Q8bhn2	mus musculu
875	5	55.6	10	2	Q9THM6_9MYRT	Q9thm6	leptospermu	948	5	55.6	10	2	Q8CIN5_CAVPO	Q8cin5	cavia porce
876	5	55.6	10	2	Q9THM7_9MYRT	Q9thm7	leptospermu	949	5	55.6	10	2	Q8CJ31_MOUSE	Q8cj31	mus musculu
877	5	55.6	10	2	Q9TKE0_9MYRT	Q9tke0	neofabricia	950	5	55.6	10	2	Q91WZ3_9MURI	Q91wz3	rattus sp.
878	5	55.6	10	2	Q9TKE1_9MYRT	Q9tke1	neofabricia	951	5	55.6	10	2	Q9QVE5_9MURI	Q9qve5	mus sp. pro
879	5	55.6	10	2	Q9TKE2_9MYRT	Q9tke2	leptospermu	952	5	55.6	10	2	Q9QVE6_9MURI	Q9qve6	mus sp. pro
880	5	55.6	10	2	Q9TKE3_9MYRT	Q9tke3	leptospermu	953	5	55.6	10	2	Q9QVE7_9MURI	Q9qve7	mus sp. pro
881	5	55.6	10	2	Q9TKE4_9MYRT	Q9tke4	leptospermu	954	5	55.6	10	2	Q9QVE8_9MURI	Q9qve8	mus sp. pro
882	5	55.6	10	2	Q9TKE6_9MYRT	Q9tke6	leptospermu	955	5	55.6	10	2	Q9QVF0_9MURI	Q9qvf0	mus sp. pro
883	5	55.6	10	2	Q9TKE7_KUNPU	Q9tke7	kunzea pulc	956	5	55.6	10	2	Q9QVF1_9MURI	Q9qvf1	mus sp. pro
884	5	55.6	10	2	Q9TKE8_KUNER	Q9tke8	kunzea eric	957	5	55.6	10	2	Q9QVF7_9MURI	Q9qvf7	rattus sp.
885	5	55.6	10	2	Q9TKE9_KUNBA	Q9tke9	kunzea baxt	958	5	55.6	10	2	Q9T2P3_9MURI	Q9t2p3	rattus sp.
886	5	55.6	10	2	Q9TKF0_KUNAM	Q9tkf0	kunzea ambi	959	5	55.6	10	2	Q9Z2V3_RAT	Q9z2v3	rattus norv
887	5	55.6	10	2	Q9TKF1_9MYRT	Q9tkf1	homalosperm	960	5	55.6	10	2	Q53VQ3_MOUSE	Q53vq3	mus musculu
888	5	55.6	10	2	Q9TKF3_9MYRT	Q9tkf3	asteromyrtu	961	5	55.6	10	2	Q53VQ7_MOUSE	Q53vq7	mus musculu
889	5	55.6	10	2	Q9TKF4_9MYRT	Q9tkf4	angasomyrtu	962	5	55.6	10	2	Q53VR1_MOUSE	Q53vr1	mus musculu
890	5	55.6	10	2	Q9TKF5_9MYRT	Q9tkf5	agonis spat	963	5	55.6	10	2	Q7TSC5_MOUSE	Q7tsc5	mus musculu
891	5	55.6	10	2	Q9TKF6_9MYRT	Q9tkf6	agonis obtu	964	5	55.6	10	2	Q5TLE7_MOUSE	Q5tle7	mus musculu
892	5	55.6	10	2	Q9TKF7_9MYRT	Q9tkf7	agonis gran	965	5	55.6	10	2	Q6LBT3_MOUSE	Q6lbt3	mus musculu
893	5	55.6	10	2	Q9TKF8_9MYRT	Q9tkf8	tristaniops	966	5	55.6	10	2	Q80WD3_MUSSP	Q80wd3	mus spretus
894	5	55.6	10	2	Q9TKF9_MELVI	Q9tkf9	melaleuca v	967	5	55.6	10	2	O90346_9FLAV	O90346	gb virus c/
895	5	55.6	10	2	Q9TKG0_9MYRT	Q9tkg0	lophostemon	968	5	55.6	10	2	O90347_9FLAV	O90347	gb virus c/
896	5	55.6	10	2	Q9TKG2_CALPO	Q9tkg2	callistemon	969	5	55.6	10	2	O90348_9FLAV	O90348	gb virus c/
897	5	55.6	10	2	Q9XMB4_AEGTA	Q9xmb4	aegilops ta	970	5	55.6	10	2	O90349_9FLAV	O90349	gb virus c/
898	5	55.6	10	2	Q56ZK9_ARATH	Q56zk9	arabidopsis	971	5	55.6	10	2	Q64971_9BROM	Q64971	alfalfa mos
899	5	55.6	10	2	Q5SBS2_9POAL	Q5sbs2	thamnochort	972	5	55.6	10	2	Q66M69_9VIRU	Q66m69	potato viru
900	5	55.6	10	2	P82132_SPIOL	P82132	spinacia ol	973	5	55.6	10	2	Q69347_HHV1	Q69347	human herpe
901	5	55.6	10	2	P82133_SPIOL	P82133	spinacia ol	974	5	55.6	10	2	Q70GQ6_PRVKA	Q70gq6	pseudorabie
902	5	55.6	10	2	P82136_SPIOL	P82136	spinacia ol	975	5	55.6	10	2	Q76V79_9POLY	Q76v79	polyomaviru
903	5	55.6	10	2	Q4QWV8_9MARC	Q4qwv8	jensenia di	976	5	55.6	10	2	Q53X10_POVBK	Q53x10	polyomaviru
904	5	55.6	10	2	Q44693_BACAM	Q44693	bacillus am	977	5	55.6	10	2	Q53X11_POVBK	Q53x11	polyomaviru
905	5	55.6	10	2	Q50032_MYCLE	Q50032	mycobacteri	978	5	55.6	10	2	Q53X32_POVBK	Q53x32	polyomaviru
906	5	55.6	10	2	Q52837_RHILE	Q52837	rhizobium l	979	5	55.6	10	2	Q53X33_POVBK	Q53x33	polyomaviru
907	5	55.6	10	2	Q5D4Q3_9RHOO	Q5d4q3	azospira or	980	5	55.6	10	2	Q82625_IBDV	Q82625	avian infec

981	5	55.6	10	2	O42355_BRARE	O42355 brachydanio
982	5	55.6	10	2	O73589_CHICK	O73589 gallus gall
983	5	55.6	10	2	O73594_CHICK	O73594 gallus gall
984	5	55.6	10	2	O79885_9SAUR	O79885 anolis pate
985	5	55.6	10	2	O79888_BASPL	O79888 basiliscus
986	5	55.6	10	2	O79891_CROCO	O79891 crotaphytus
987	5	55.6	10	2	O79894_GAMWI	O79894 gambelia wi
988	5	55.6	10	2	O79897_HOPSP	O79897 hoplocercus
989	5	55.6	10	2	O79900_9SAUR	O79900 liolaemus p
990	5	55.6	10	2	O79903_OPLCU	O79903 oplurus cuv
991	5	55.6	10	2	O79906_9SAUR	O79906 phrynosoma
992	5	55.6	10	2	O79909_SAUAT	O79909 sauromalus
993	5	55.6	10	2	O79912_CHAFI	O79912 chamaeleo f
994	5	55.6	10	2	O79915_LEIBE	O79915 leiolepis b
995	5	55.6	10	2	O79924_9SAUR	O79924 elgaria pan
996	5	55.6	10	2	P92576_BIPBI	P92576 bipes bipor
997	5	55.6	10	2	P92616_9SAUR	P92616 aspidosceli
998	5	55.6	10	2	P92648_9SAUR	P92648 lialis jica
999	5	55.6	10	2	P92758_9SAUR	P92758 teratoscinc
1000	5	55.6	10	2	P92762_UROAC	P92762 uromastyx a

ALIGNMENTS

```
RESULT 1
FAR3_HIRME
ID   FAR3_HIRME     STANDARD;      PRT;     4 AA.
AC   P42562;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide YLRF-amide.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933; DOI=10.1016/0196-9781(91)90035-N;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       4       4       Phenylalanine amide.
SQ   SEQUENCE   4 AA;  598 MW;  69D4073B30000000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 2
FAR4_HIRME
ID   FAR4_HIRME     STANDARD;      PRT;     4 AA.
AC   P42563;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide YMRF-amide.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933; DOI=10.1016/0196-9781(91)90035-N;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       4       4       Phenylalanine amide.
SQ   SEQUENCE   4 AA;  616 MW;  69D4068B30000000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 3
FLRF_HELTI
ID   FLRF_HELTI     STANDARD;      PRT;     4 AA.
AC   P69138; P42561;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   FLRFamide.
OS   Helisoma trivolvis (Snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC   Lymnaeoidea; Planorbidae; Helisoma.
OX   NCBI_TaxID=27815;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=94286417; PubMed=7912428; DOI=10.1016/0196-9781(94)90166-X;
RA   Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT   "FMRFamide-related peptides from the kidney of the snail, Helisoma
RT   trivolvis.";
RL   Peptides 15:31-36(1994).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       4       4       Phenylalanine amide.
SQ   SEQUENCE   4 AA;  582 MW;  69D40729A0000000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
```

```
                |
Db          3 R 3


RESULT 4
FLRF_HIRME
ID   FLRF_HIRME     STANDARD;      PRT;     4 AA.
AC   P69137; P42561;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   FLRFamide.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933; DOI=10.1016/0196-9781(91)90035-N;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       4       4       Phenylalanine amide.
SQ   SEQUENCE   4 AA;  582 MW;  69D40729A0000000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
                |
Db          3 R 3


RESULT 5
FLRN_ANTEL
ID   FLRN_ANTEL     STANDARD;      PRT;     4 AA.
AC   P58707;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Antho-RNamide.
OS   Anthopleura elegantissima (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Anthopleura.
OX   NCBI_TaxID=6110;
RN   [1]
RP   PROTEIN SEQUENCE, AND MASS SPECTROMETRY.
RX   MEDLINE=90319122; PubMed=1973541;
RA   Grimmelikhuijzen C.J.P., Rinehart K.L. Jr., Jacob E., Graff D.,
RA   Reinscheid R.K., Nothacker H.-P., Staley A.L.;
RT   "Isolation of L-3-phenyllactyl-Leu-Arg-Asn-NH2 (Antho-RNamide), a sea
RT   anemone neuropeptide containing an unusual amino-terminal blocking
RT   group.";
RL   Proc. Natl. Acad. Sci. U.S.A. 87:5410-5414(1990).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Neuron specific.
CC   -!- MASS SPECTROMETRY: MW=549.3; METHOD=FAB; RANGE=1-4; NOTE=Ref.1.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A35779; A35779.
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       1       1       3-phenyllactic acid.
FT   MOD_RES       4       4       Asparagine amide.
SQ   SEQUENCE   4 AA;  549 MW;  64540729A0000000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
                |
Db          3 R 3


RESULT 6
FMRF_HELTI
ID   FMRF_HELTI     STANDARD;      PRT;     4 AA.
AC   P69148; P01162;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   FMRFamide.
OS   Helisoma trivolvis (Snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC   Lymnaeoidea; Planorbidae; Helisoma.
OX   NCBI_TaxID=27815;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=94286417; PubMed=7912428; DOI=10.1016/0196-9781(94)90166-X;
RA   Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT   "FMRFamide-related peptides from the kidney of the snail, Helisoma
RT   trivolvis.";
RL   Peptides 15:31-36(1994).
CC   -!- FUNCTION: Myoactive; cardioexcitatory substance. Pharmacological
CC       activities include augmentation, induction, and regularization of
CC       cardiac contraction.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       4       4       Phenylalanine amide.
SQ   SEQUENCE   4 AA;  600 MW;  69D40699A0000000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
                |
Db          3 R 3


RESULT 7
FMRF_HIRME
ID   FMRF_HIRME     STANDARD;      PRT;     4 AA.
AC   P69147; P01162;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   FMRFamide.
OS   Hirudo medicinalis (Medicinal leech).
```

```
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933; DOI=10.1016/0196-9781(91)90035-N;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- FUNCTION: Myoactive; cardioexcitatory substance. Pharmacological
CC       activities include augmentation, induction, and regularization of
CC       cardiac contraction.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       4       4       Phenylalanine amide.
SQ   SEQUENCE   4 AA;  600 MW;  69D40699A0000000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          3 R 3


RESULT 8
FMRF_MACNI
ID   FMRF_MACNI     STANDARD;      PRT;     4 AA.
AC   P69145; P01162;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   FMRFamide (Peak C) (Cardioexcitatory neuropeptide).
OS   Macrocallista nimbosa (Sun-ray clam).
OC   Eukaryota; Metazoa; Mollusca; Bivalvia; Heteroconchia; Veneroida;
OC   Veneroidea; Veneridae; Macrocallista.
OX   NCBI_TaxID=6594;
RN   [1]
RP   PROTEIN SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Cerebral pedal, and Visceral ganglion;
RX   MEDLINE=77215956; PubMed=877582;
RA   Price D.A., Greenberg M.J.;
RT   "Structure of a molluscan cardioexcitatory neuropeptide.";
RL   Science 197:670-671(1977).
RN   [2]
RP   PROTEIN SEQUENCE, AND CHARACTERIZATION.
RC   TISSUE=Ganglion;
RX   MEDLINE=78012038; PubMed=909875;
RA   Price D.A., Greenberg M.J.;
RT   "Purification and characterization of a cardioexcitatory neuropeptide
RT   from the central ganglia of a bivalve mollusc.";
RL   Prep. Biochem. 7:261-281(1977).
CC   -!- FUNCTION: Myoactive; cardioexcitatory substance. Pharmacological
CC       activities include augmentation, induction, and regularization of
CC       cardiac contraction.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A01426; ECNK.
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       4       4       Phenylalanine amide.
SQ   SEQUENCE   4 AA;  600 MW;  69D40699A0000000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          3 R 3


RESULT 9
FMRF_NERVI
ID   FMRF_NERVI     STANDARD;      PRT;     4 AA.
AC   P69146; P01162;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   FMRFamide.
OS   Nereis virens (Sandworm).
OC   Eukaryota; Metazoa; Annelida; Polychaeta; Palpata; Aciculata;
OC   Phyllodocida; Nereididae; Nereis.
OX   NCBI_TaxID=6353;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=90259866; PubMed=2342992; DOI=10.1016/0196-9781(90)90113-J;
RA   Krajniak K.G., Price D.A.;
RT   "Authentic FMRFamide is present in the polychaete Nereis virens.";
RL   Peptides 11:75-77(1990).
CC   -!- FUNCTION: Myoactive; cardioexcitatory substance. Pharmacological
CC       activities include augmentation, induction, and regularization of
CC       cardiac contraction.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A60418; A60418.
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       4       4       Phenylalanine amide.
SQ   SEQUENCE   4 AA;  600 MW;  69D40699A0000000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          3 R 3


RESULT 10
FYRI_ANTEL
ID   FYRI_ANTEL     STANDARD;      PRT;     4 AA.
AC   P58706;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   01-FEB-2005 (Rel. 46, Last annotation update)
DE   Antho-RIamide I [Contains: Antho-RIamide II].
OS   Anthopleura elegantissima (Sea anemone).
OC   Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
OC   Nynantheae; Actiniidae; Anthopleura.
OX   NCBI_TaxID=6110;
```

```
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=92270459; PubMed=1821096; DOI=10.1016/0196-9781(91)90190-Z;
RA   Nothacker H.-P., Rinehart K.L. Jr., McFarlane I.D.,
RA   Grimmelikhuijzen C.J.P.;
RT   "Isolation of two novel neuropeptides from sea anemones: the unusual,
RT   biologically active L-3-phenyllactyl-Tyr-Arg-Ile-NH2 and its des-
RT   phenyllactyl fragment Tyr-Arg-Ile-NH2.";
RL   Peptides 12:1165-1173(1991).
RN   [2]
RP   FUNCTION.
RX   MEDLINE=93391436; PubMed=8397415;
RA   McFarlane I.D., Hudman D., Nothacker H.-P., Grimmelikhuijzen C.J.P.;
RT   "The expansion behaviour of sea anemones may be coordinated by two
RT   inhibitory neuropeptides, Antho-KAamide and Antho-RIamide.";
RL   Proc. R. Soc. Lond., B, Biol. Sci. 253:183-188(1993).
CC   -!- FUNCTION: Inhibits spontaneous contractions in several muscle
CC       groups. May be involved in the expansion phase of feeding
CC       behaviour in sea anemones.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Neuron specific.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   PEPTIDE       1      4       Antho-RIamide I.
FT   PEPTIDE       2      4       Antho-RIamide II.
FT   MOD_RES       1      1       3-phenyllactic acid.
FT   MOD_RES       4      4       Isoleucine amide.
SQ   SEQUENCE   4 AA;  598 MW;  60441B59A0000000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 11
TUFT_HUMAN
ID   TUFT_HUMAN     STANDARD;      PRT;     4 AA.
AC   P01858;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Phagocytosis-stimulating peptide (Tuftsin).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=72187087; PubMed=4112769;
RA   Nishioka K., Constantopoulos A., Satoh P.S., Najjar V.A.;
RT   "The characteristics, isolation and synthesis of the phagocytosis
RT   stimulating peptide tuftsin.";
RL   Biochem. Biophys. Res. Commun. 47:172-179(1972).
RN   [2]
RP   IMMUNOGLOBULIN CLASS.
RX   MEDLINE=68091045; PubMed=4169272;
RA   Fidalgo B.V., Najjar V.A.;
RT   "The physiological role of the lymphoid system. VI. The stimulatory
RT   effect of leucophilic gamma globulin (leucokinin) on the phagocytic
RT   activity of human polymorphonuclear leucocyte.";
RL   Biochemistry 6:3386-3392(1967).
CC   -!- MISCELLANEOUS: An IgG (called leucokinin) binds reversibly to the
CC       cell membrane of neutrophils in the blood. Leucokininase on the
CC       membrane releases the active peptide tuftsin from the gamma chain.
CC       Tuftsin is essential for maximum stimulation of the phagocytic
CC       activity of neutrophils.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A02147; A02147.
DR   MIM; 191150; -.
DR   GO; GO:0003823; F:antigen binding; NAS.
DR   GO; GO:0006909; P:phagocytosis; NAS.
KW   Direct protein sequencing.
SQ   SEQUENCE   4 AA;  501 MW;  74176321C0000000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 4;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          4 R 4


RESULT 12
FARP_ARTTR
ID   FARP_ARTTR     STANDARD;      PRT;     5 AA.
AC   P41853;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide RYIRF-amide.
OS   Artioposthia triangulata (New Zealand flatworm).
OC   Eukaryota; Metazoa; Platyhelminthes; Turbellaria; Seriata; Tricladida;
OC   Terricola; Geoplanidae; Arthurdendyus.
OX   NCBI_TaxID=132421;
RN   [1]
RP   PROTEIN SEQUENCE, AND SYNTHESIS.
RX   MEDLINE=94211927; PubMed=7909164; DOI=10.1016/0167-0115(94)90189-9;
RA   Maule A.G., Shaw C., Halton D.W., Curry W.J., Thim L.;
RT   "RYIRFamide: a turbellarian FMRFamide-related peptide (FaRP).";
RL   Regul. Pept. 50:37-43(1994).
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       5      5       Phenylalanine amide.
SQ   SEQUENCE   5 AA;  754 MW;  69D4004B44600000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 13
FARP_CHICK
ID   FARP_CHICK     STANDARD;      PRT;     5 AA.
AC   P83308;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
```

```
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide (LPLRF-amide).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   PROTEIN SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Brain;
RX   PubMed=6137771;
RA   Dockray G.J., Reeve J.R. Jr., Shively J., Gayton R.J., Barnard C.S.;
RT   "A novel active pentapeptide from chicken brain identified by
RT   antibodies to FMRFamide.";
RL   Nature 305:328-330(1983).
CC   -!- FUNCTION: May function as a neurotransmitter or modulator.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   GO; GO:0007218; P:neuropeptide signaling pathway; TAS.
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       5       5       Phenylalanine amide.
SQ   SEQUENCE   5 AA;  645 MW;  69D4073767400000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          4 R 4


RESULT 14
PRCT_CARMA
ID   PRCT_CARMA     STANDARD;      PRT;     5 AA.
AC   P67857; P01373;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Proctolin.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=86232789; PubMed=2872661; DOI=10.1016/0196-9781(86)90063-X;
RA   Stangier J., Dircksen H., Keller R.;
RT   "Identification and immunocytochemical localization of proctolin in
RT   pericardial organs of the shore crab, Carcinus maenas.";
RL   Peptides 7:67-72(1986).
CC   -!- FUNCTION: Stimulates cardiac output and hindgut motility,
CC       modulates visceral and skeletal muscle in many arthropods.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Found in the crab pericardial organs.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Direct protein sequencing; Neuropeptide.
SQ   SEQUENCE   5 AA;  649 MW;  71B7673B44600000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 15
PRCT_LIMPO
ID   PRCT_LIMPO     STANDARD;      PRT;     5 AA.
AC   P67858; P01373;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Proctolin.
OS   Limulus polyphemus (Atlantic horseshoe crab).
OC   Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC   Limulidae; Limulus.
OX   NCBI_TaxID=6850;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=90287800; PubMed=2356151; DOI=10.1016/0196-9781(90)90072-D;
RA   Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,
RA   Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,
RA   Shabanowitz J.;
RT   "Identification of proctolin in the central nervous system of the
RT   horseshoe crab, Limulus polyphemus.";
RL   Peptides 11:205-211(1990).
CC   -!- FUNCTION: Stimulates cardiac output and hindgut motility,
CC       modulates visceral and skeletal muscle in many arthropods.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Found in the crab pericardial organs.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A60411; A60411.
KW   Direct protein sequencing; Neuropeptide.
SQ   SEQUENCE   5 AA;  649 MW;  71B7673B44600000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 16
PRCT_PERAM
ID   PRCT_PERAM     STANDARD;      PRT;     5 AA.
AC   P67859; P01373;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   Proctolin.
OS   Periplaneta americana (American cockroach).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattoidea;
OC   Blattidae; Blattinae; Periplaneta.
OX   NCBI_TaxID=6978;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=76074708; PubMed=576; DOI=10.1016/0024-3205(75)90134-4;
RA   Starratt A.N., Brown B.E.;
RT   "Structure of the pentapeptide proctolin, a proposed neurotransmitter
```

```
RT   in insects.";
RL   Life Sci. 17:1253-1256(1975).
RN   [2]
RP   BIOLOGICAL SOURCE.
RX   MEDLINE=81225865; PubMed=6113690;
RA   O'Shea M., Adams M.E.;
RT   "Pentapeptide (proctolin) associated with an identified neuron.";
RL   Science 213:567-569(1981).
CC   -!- FUNCTION: Stimulates cardiac output and hindgut motility,
CC       modulates visceral and skeletal muscle in many arthropods.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Found in the lateral white neurons.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A01644; HOROHA.
KW   Direct protein sequencing; Neuropeptide.
SQ   SEQUENCE   5 AA;  649 MW;  71B7673B44600000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          1 R 1


RESULT 17
UF01_MOUSE
ID   UF01_MOUSE     STANDARD;      PRT;     5 AA.
AC   P38639;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Unknown protein from 2D-PAGE of fibroblasts (P19) (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   TISSUE=Fibroblast;
RX   MEDLINE=95009907; PubMed=7523108;
RA   Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT   "Separation and sequencing of familiar and novel murine proteins using
RT   preparative two-dimensional gel electrophoresis.";
RL   Electrophoresis 15:735-745(1994).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 6.6, its MW is: 19 kDa.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Direct protein sequencing.
FT   NON_TER       5      5
SQ   SEQUENCE   5 AA;  717 MW;  7364087043100000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 5;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          4 R 4


RESULT 18
ACPH_RABIT
ID   ACPH_RABIT     STANDARD;      PRT;     6 AA.
AC   P25154;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   01-FEB-2005 (Rel. 46, Last annotation update)
DE   Acylamino-acid-releasing enzyme (EC 3.4.19.1) (AARE) (Acyl-peptide
DE   hydrolase) (APH) (Acylaminoacyl-peptidase) (Fragment).
GN   Name=APEH;
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Lagomorpha; Leporidae;
OC   Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   TISSUE=Muscle;
RX   MEDLINE=92222120; PubMed=1807161;
RA   Krishna R.G., Chin C.C.Q., Wold F.;
RT   "N-terminal sequence analysis of N alpha-acetylated proteins after
RT   unblocking with N-acylaminoacyl-peptide hydrolase.";
RL   Anal. Biochem. 199:45-50(1991).
CC   -!- FUNCTION: This enzyme catalyzes the hydrolysis of the N-terminal
CC       peptide bond of an N-acetylated peptide to generate an N-
CC       acetylated amino acid and a peptide with a free N-terminus. It
CC       preferentially cleaves off Ac-Ala, Ac-Met and Ac-Ser.
CC   -!- CATALYTIC ACTIVITY: Cleavage of an N-acetyl or N-formyl amino acid
CC       from the N-terminus of a polypeptide.
CC   -!- SUBUNIT: Homotetramer.
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- SIMILARITY: Belongs to the peptidase S9C family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A49792; A49792.
DR   InterPro; IPR002471; Pept_S9_AS.
DR   PROSITE; PS00708; PRO_ENDOPEP_SER; PARTIAL.
KW   Acetylation; Direct protein sequencing; Hydrolase.
FT   MOD_RES       1      1       N-acetylmethionine.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  775 MW;  6732D6C40B16F000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


RESULT 19
FARP_MONEX
ID   FARP_MONEX     STANDARD;      PRT;     6 AA.
AC   P41966;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide GNFFRF-amide.
OS   Moniezia expansa (Sheep tapeworm).
OC   Eukaryota; Metazoa; Platyhelminthes; Cestoda; Eucestoda;
OC   Cyclophyllidea; Anoplocephalidae; Moniezia.
OX   NCBI_TaxID=28841;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=93312289; PubMed=8323531;
```

```
RA   Maule A.G., Shaw C., Halton D.W., Thim L.;
RT   "GNFFRFamide: a novel FMRFamide-immunoreactive peptide isolated from
RT   the sheep tapeworm, Moniezia expansa.";
RL   Biochem. Biophys. Res. Commun. 193:1054-1060(1993).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A43129; A43129.
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       6      6       Phenylalanine amide.
SQ   SEQUENCE   6 AA;  787 MW;  69D409C9C4481000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          5 R 5


RESULT 20
PYF1_PENMO
ID   PYF1_PENMO     STANDARD;      PRT;     6 AA.
AC   P84005;
DT   05-JUL-2004 (Rel. 44, Created)
DT   05-JUL-2004 (Rel. 44, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Peptide tyrosine phenylalanine 1 (Pem-PYF1).
OS   Penaeus monodon (Penoeid shrimp).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC   Penaeidae; Penaeus.
OX   NCBI_TaxID=6687;
RN   [1]
RP   PROTEIN SEQUENCE, TISSUE SPECIFICITY, AND MASS SPECTROMETRY.
RC   TISSUE=Eyestalk;
RX   PubMed=12431727; DOI=10.1016/S0196-9781(02)00176-6;
RA   Sithigorngul P., Pupuem J., Krungkasem C., Longyant S., Panchan N.,
RA   Chaivisuthangkura P., Sithigorngul W., Petsom A.;
RT   "Four novel PYFs: members of NPY/PP peptide superfamily from the
RT   eyestalk of the giant tiger prawn Penaeus monodon.";
RL   Peptides 23:1895-1906(2002).
CC   -!- FUNCTION: May act as a neurotransmitter, neuromodulator or
CC       neurohormone.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Limited to neuronal cell bodies, neuronal
CC       processes and sinus gland.
CC   -!- MASS SPECTROMETRY: MW=801.5; METHOD=MALDI; RANGE=1-6; NOTE=Ref.1.
CC   -!- SIMILARITY: Belongs to the NPY family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   InterPro; IPR001955; Pancreatic_hormn.
DR   PROSITE; PS00265; PANCREATIC_HORMONE_1; PARTIAL.
DR   PROSITE; PS50276; PANCREATIC_HORMONE_2; PARTIAL.
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       6      6       Phenylalanine amide (Potential).
SQ   SEQUENCE   6 AA;  802 MW;  69D417740DC46000 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          1 R 1


RESULT 21
P82181_SPIOL
ID   P82181_SPIOL PRELIMINARY;      PRT;     6 AA.
AC   P82181;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 50S ribosomal protein L10 beta (Fragment).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435798; PubMed=10874046; DOI=10.1074/jbc.M005012200;
RA   Yamaguchi K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the 50 S subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 275:28466-28482(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 kDa.
CC   -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002363; Ribosomal_L10eub.
DR   PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW   Chloroplast; Ribosomal protein; rRNA-binding.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  675 MW;  6321B415B05DB000 CRC64;

  Query Match             55.6%;  Score 5;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          4 R 4


RESULT 22
P82182_SPIOL
ID   P82182_SPIOL PRELIMINARY;      PRT;     6 AA.
AC   P82182;
DT   01-JUN-2000 (TrEMBLrel. 14, Created)
DT   01-JUN-2000 (TrEMBLrel. 14, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 50S ribosomal protein L10 gamma (Fragment).
OS   Spinacia oleracea (Spinach).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435798; PubMed=10874046; DOI=10.1074/jbc.M005012200;
RA   Yamaguchi K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the 50 S subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 275:28466-28482(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA.
```

```
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 16.5 kDa.
CC   -!- SIMILARITY: BELONGS TO THE L10P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002363; Ribosomal_L10eub.
DR   PROSITE; PS01109; RIBOSOMAL_L10; PARTIAL.
KW   Chloroplast; Ribosomal protein; rRNA-binding.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  675 MW;  6321B415B05DB000 CRC64;

  Query Match             55.6%;  Score 5;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          4 R 4


RESULT 23
P82541_SPIOL
ID   P82541_SPIOL PRELIMINARY;      PRT;     6 AA.
AC   P82541;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Chloroplast 30S ribosomal protein S19 beta (Fragment).
OS   Spinacia oleracea (Spinach).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   PROTEIN SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC   STRAIN=cv. ALWARO; TISSUE=Leaf;
RX   MEDLINE=20435797; PubMed=10874039; DOI=10.1074/jbc.M004350200;
RA   Yamaguchi K., von Knoblauch K., Subramanian A.R.;
RT   "The plastid ribosomal proteins. Identification of all the proteins in
RT   the 30S subunit of an organelle ribosome (chloroplast).";
RL   J. Biol. Chem. 275:28455-28465(2000).
CC   -!- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 16S RIBOSOMAL RNA.
CC   -!- SUBCELLULAR LOCATION: CHLOROPLAST.
CC   -!- TISSUE SPECIFICITY: EXPRESSED IN ALL PLANT TISSUES.
CC   -!- MASS SPECTROMETRY: MW=10477.0; METHOD=ELECTROSPRAY.
CC   -!- MASS SPECTROMETRY: MW=10495; METHOD=MALDI.
CC   -!- MISCELLANEOUS: S19 ALPHA AND BETA FORMS DIFFER IN PI. S19 BETA
CC       FORM IS THE MINOR BASIC FORM.
CC   -!- MISCELLANEOUS: ON THE 2D-GEL ITS MW IS: 12 kDa.
CC   -!- SIMILARITY: BELONGS TO THE S19P FAMILY OF RIBOSOMAL PROTEINS.
DR   GO; GO:0009507; C:chloroplast; IEA.
DR   GO; GO:0019843; F:rRNA binding; IEA.
DR   GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR   InterPro; IPR002222; Ribosomal_S19.
DR   PROSITE; PS00323; RIBOSOMAL_S19; PARTIAL.
KW   Chloroplast; Ribosomal protein; rRNA-binding.
FT   NON_TER       6      6
SQ   SEQUENCE   6 AA;  732 MW;  63333735A411C000 CRC64;

  Query Match             55.6%;  Score 5;  DB 2;  Length 6;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 24
CAP6_CANAL
ID   CAP6_CANAL     STANDARD;      PRT;     7 AA.
AC   P83784;
DT   01-FEB-2005 (Rel. 46, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   01-FEB-2005 (Rel. 46, Last annotation update)
DE   Cytoplasmic antigenic protein 6 (Fragment).
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   PROTEIN SEQUENCE, SUBCELLULAR LOCATION, AND ANTIGENICITY.
RC   STRAIN=SC5314; TISSUE=Protoplast;
RX   PubMed=15378761; DOI=10.1002/pmic.200400903;
RA   Pitarch A., Abian J., Carrascal M., Sanchez M., Nombela C., Gil C.;
RT   "Proteomics-based identification of novel Candida albicans antigens
RT   for diagnosis of systemic candidiasis in patients with underlying
RT   hematological malignancies.";
RL   Proteomics 4:3084-3106(2004).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: Has antigenic properties. Elicits a specific immune
CC       response in systemic candidiasis human patients undergoing
CC       malignant hematological disorders.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Antigen; Direct protein sequencing.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  900 MW;  740736C6D046DAC0 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          7 R 7


RESULT 25
CARP_MYTED
ID   CARP_MYTED     STANDARD;      PRT;     7 AA.
AC   P10420;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Catch-relaxing peptide (CARP).
OS   Mytilus edulis (Blue mussel).
OC   Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Mytiloida;
OC   Mytiloidea; Mytilidae; Mytilinae; Mytilus.
OX   NCBI_TaxID=6550;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=88052022; PubMed=3676797;
RA   Hirata T., Kubota I., Takabatake I., Kawahara A., Shimamoto N.,
RA   Muneoka Y.;
RT   "Catch-relaxing peptide isolated from Mytilus pedal ganglia.";
RL   Brain Res. 422:374-376(1987).
CC   -!- FUNCTION: This peptide exhibits both potentiating (contraction)
CC       and inhibitory (relaxation) effects on the anterior byssus
CC       retractor muscle.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
```

```
DR   PIR; A29342; ECMUCR.
KW   Amidation; Direct protein sequencing; Hormone.
FT   MOD_RES       7      7       Leucine amide.
SQ   SEQUENCE   7 AA;  831 MW;  6734072687669DB0 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          6 R 6


RESULT 26
CHOX_ALCSP
ID   CHOX_ALCSP     STANDARD;      PRT;     7 AA.
AC   P16101;
DT   01-APR-1990 (Rel. 14, Created)
DT   01-APR-1990 (Rel. 14, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Choline oxidase (EC 1.1.3.17) (Fragment).
OS   Alcaligenes sp.
OC   Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC   Alcaligenaceae; Alcaligenes.
OX   NCBI_TaxID=512;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=81006769; PubMed=6997283;
RA   Ohta-Fukuyama M., Miyake Y., Emi S., Yamano T.;
RT   "Identification and properties of the prosthetic group of choline
RT   oxidase from Alcaligenes sp.";
RL   J. Biochem. 88:197-203(1980).
CC   -!- CATALYTIC ACTIVITY: Choline + O(2) = betaine aldehyde + H(2)O(2).
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; A15398; A15398.
KW   Direct protein sequencing; Oxidoreductase.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  839 MW;  7415B1E457644AC0 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          7 R 7


RESULT 27
FAF1_ASCSU
ID   FAF1_ASCSU     STANDARD;      PRT;     7 AA.
AC   P31889;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   FMRFamide-like neuropeptide AF1.
OS   Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC   Ascarididae; Ascaris.
OX   NCBI_TaxID=6253;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=90180465; PubMed=2627377;
RA   Cowden C., Stretton A.O.W., Davis R.E.;
RT   "AF1, a sequenced bioactive neuropeptide isolated from the nematode
RT   Ascaris suum.";
RL   Neuron 2:1465-1473(1989).
CC   -!- FUNCTION: Potent modulator of inhibitory motorneurons. Reduces the
CC       input resistance and blocks slow oscillatory potentials in these
CC       cells.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Found in the nerve cords and a variety of
CC       ganglia particularly in the anterior regions.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7      7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  953 MW;  69D40059CB144350 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          6 R 6


RESULT 28
FAF2_ASCSU
ID   FAF2_ASCSU     STANDARD;      PRT;     7 AA.
AC   P67879; P31890;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   FMRFamide-like neuropeptide AF2.
OS   Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC   Ascarididae; Ascaris.
OX   NCBI_TaxID=6253;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=93324431; PubMed=8332542; DOI=10.1016/0196-9781(93)90127-3;
RA   Cowden C., Stretton A.O.W.;
RT   "AF2, an Ascaris neuropeptide: isolation, sequence, and bioactivity.";
RL   Peptides 14:423-430(1993).
CC   -!- FUNCTION: Has effects on muscle tension.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Found in the nerve cords and a variety of
CC       ganglia particularly in the anterior regions.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7      7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  992 MW;  69D4073B5B11E350 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          6 R 6
```

```
RESULT 29
FAF2_PANRE
ID   FAF2_PANRE     STANDARD;      PRT;     7 AA.
AC   P67880; P31890;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   25-OCT-2004 (Rel. 45, Last annotation update)
DE   FMRFamide-like neuropeptide AF2.
OS   Panagrellus redivivus.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC   Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX   NCBI_TaxID=6233;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=95060998; PubMed=7970891;
RA   Maule A.G., Shaw C., Bowman J.W.;
RT   "The FMRFamide-like neuropeptide AF2 (Ascaris suum) is present in the
RT   free-living nematode, Panagrellus redivivus (Nematoda, Rhabditida).";
RL   Parasitology 109:351-356(1994).
CC   -!- FUNCTION: Has effects on muscle tension.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Found in the nerve cords and a variety of
CC       ganglia particularly in the anterior regions.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7      7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  992 MW;  69D4073B5B11E350 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          6 R 6


RESULT 30
FAR1_HELTI
ID   FAR1_HELTI     STANDARD;      PRT;     7 AA.
AC   P41871;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide GDPFLRF-amide.
OS   Helisoma trivolvis (Snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Basommatophora;
OC   Lymnaeoidea; Planorbidae; Helisoma.
OX   NCBI_TaxID=27815;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   TISSUE=Kidney;
RX   MEDLINE=94286417; PubMed=7912428; DOI=10.1016/0196-9781(94)90166-X;
RA   Madrid K.P., Price D.A., Greenberg M.J., Khan H.R., Saleuddin A.S.M.;
RT   "FMRFamide-related peptides from the kidney of the snail, Helisoma
RT   trivolvis.";
RL   Peptides 15:31-36(1994).
CC   -!- FUNCTION: Appears to be involved in osmoregulation by affecting
CC       the kidney, mantle and skin.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Kidney, skin, mantle and the hemolymph.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7      7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  851 MW;  69D40729D76AA810 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          6 R 6


RESULT 31
FAR1_MACRS
ID   FAR1_MACRS     STANDARD;      PRT;     7 AA.
AC   P83274;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide FLP1 (DRNFLRF-amide).
OS   Macrobrachium rosenbergii (Giant fresh water prawn).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Caridea;
OC   Palaemonoidea; Palaemonidae; Macrobrachium.
OX   NCBI_TaxID=79674;
RN   [1]
RP   PROTEIN SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Eyestalk;
RA   Sithigorngul P., Saraithongkum W., Jaideechoey S., Longyant S.,
RA   Sithigorngul W.;
RT   "Novel FMRFamide-like neuropeptides from the eyestalk of the giant
RT   freshwater prawn Macrobrachium rosenbergii.";
RL   Comp. Biochem. Physiol. 120B:587-595(1998).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- MASS SPECTROMETRY: MW=965.7; METHOD=MALDI; RANGE=1-7; NOTE=Ref.1.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   GO; GO:0007218; P:neuropeptide signaling pathway; TAS.
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7      7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  967 MW;  69D40729C4540AC0 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          2 R 2


RESULT 32
FAR1_PROCL
ID   FAR1_PROCL     STANDARD;      PRT;     7 AA.
AC   P38499;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Cardioexcitatory FMRFamide homolog NF1.
OS   Procambarus clarkii (Red swamp crayfish).
```

```
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC   Astacoidea; Cambaridae; Procambarus.
OX   NCBI_TaxID=6728;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   TISSUE=Pericardial organs;
RX   MEDLINE=93248032; PubMed=8387183; DOI=10.1016/0196-9781(93)90021-8;
RA   Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;
RT   "Isolation of two FMRFamide-related peptides from crayfish pericardial
RT   organs.";
RL   Peptides 14:137-143(1993).
CC   -!- FUNCTION: Increases the rate and amplitude of spontaneous
CC       contractions of semi-isolated hearts. Increases the amplitude of
CC       excitatory postsynaptic potentials in abdominal extensor muscle.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7       7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  966 MW;  69D40729C4540420 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 33
FAR2_PROCL
ID   FAR2_PROCL     STANDARD;      PRT;     7 AA.
AC   P38498;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Cardioexcitatory FMRFamide homolog DF2.
OS   Procambarus clarkii (Red swamp crayfish).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Astacidea;
OC   Astacoidea; Cambaridae; Procambarus.
OX   NCBI_TaxID=6728;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   TISSUE=Pericardial organs;
RX   MEDLINE=93248032; PubMed=8387183; DOI=10.1016/0196-9781(93)90021-8;
RA   Mercier A.J., Orchard I., Tebrugge V., Skerrett M.;
RT   "Isolation of two FMRFamide-related peptides from crayfish pericardial
RT   organs.";
RL   Peptides 14:137-143(1993).
CC   -!- FUNCTION: Increases the rate and amplitude of spontaneous
CC       contractions of semi-isolated hearts. Increases the amplitude of
CC       excitatory postsynaptic potentials in abdominal extensor muscle.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7       7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  967 MW;  69D40729C4540AC0 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 34
FAR3_HAECO
ID   FAR3_HAECO     STANDARD;      PRT;     7 AA.
AC   P81298;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide PF3 (KSAYMRF-amide).
OS   Haemonchus contortus (Barber pole worm).
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
OC   Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
OX   NCBI_TaxID=6289;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   TISSUE=Neuron;
RX   MEDLINE=99318264; PubMed=10391380; DOI=10.1016/S0166-6851(99)00057-2;
RA   Marks N.J., Sangster N.C., Maule A.G., Halton D.W., Thompson D.P.,
RA   Geary T.G., Shaw C.;
RT   "Structural characterisation and pharmacology of KHEYLRFamide (AF2)
RT   and KSAYMRFamide (PF3/AF8) from Haemonchus contortus.";
RL   Mol. Biochem. Parasitol. 100:185-194(1999).
CC   -!- FUNCTION: Active on neuromusculature.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7       7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  902 MW;  69D4068B5DC5B350 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          6 R 6


RESULT 35
FAR3_PANRE
ID   FAR3_PANRE     STANDARD;      PRT;     7 AA.
AC   P41874;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide PF3 (KSAYMRF-amide).
OS   Panagrellus redivivus.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC   Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX   NCBI_TaxID=6233;
RN   [1]
RP   PROTEIN SEQUENCE, AND SYNTHESIS.
RX   MEDLINE=94235053; PubMed=8179635;
RA   Maule A.G., Shaw C., Bowman J.W., Halton D.W., Thompson D.P.,
```

```
RA   Geary T.G., Thim L.;
RT   "KSAYMRFamide: a novel FMRFamide-related heptapeptide from the free-
RT   living nematode, Panagrellus redivivus, which is myoactive in the
RT   parasitic nematode, Ascaris suum.";
RL   Biochem. Biophys. Res. Commun. 200:973-980(1994).
CC   -!- FUNCTION: Myoactive; induces a rapid concentration-dependent
CC       muscle tension increase.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; PC2132; PC2132.
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7       7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  902 MW;  69D4068B5DC5B350 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          6 R 6


RESULT 36
FAR4_PANRE
ID   FAR4_PANRE     STANDARD;      PRT;     7 AA.
AC   P41875;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide PF4 (KPNFIRF-amide).
OS   Panagrellus redivivus.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
OC   Panagrolaimoidea; Panagrolaimidae; Panagrellus.
OX   NCBI_TaxID=6233;
RN   [1]
RP   PROTEIN SEQUENCE, AND SYNTHESIS.
RX   MEDLINE=95232026; PubMed=7716079; DOI=10.1016/0196-9781(94)00162-Y;
RA   Maule A.G., Shaw C., Bowman J.W., Halton D.W., Thompson D.P., Thim L.,
RA   Kubiak T.M., Martin R.A., Geary T.G.;
RT   "Isolation and preliminary biological characterization of
RT   KPNFIRFamide, a novel FMRFamide-related peptide from the free-living
RT   nematode, Panagrellus redivivus.";
RL   Peptides 16:87-93(1995).
CC   -!- FUNCTION: Myoactive; induces a rapid concentration-dependent
CC       muscle tension increase.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7       7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  921 MW;  69D40059C4576350 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          6 R 6


RESULT 37
FAR5_HIRME
ID   FAR5_HIRME     STANDARD;      PRT;     7 AA.
AC   P42564;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   FMRFamide-like neuropeptide GGKYMRF-amide.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arhynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933; DOI=10.1016/0196-9781(91)90035-N;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7       7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  858 MW;  69D4068B53387810 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          6 R 6


RESULT 38
FARB_CALVO
ID   FARB_CALVO     STANDARD;      PRT;     7 AA.
AC   P41866;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   CalliFMRFamide 11.
OS   Calliphora vomitoria (Blue blowfly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Oestroidea;
OC   Calliphoridae; Calliphora.
OX   NCBI_TaxID=27454;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   TISSUE=Thoracic ganglion;
RX   MEDLINE=92196111; PubMed=1549595;
RA   Duve H., Johnsen A.H., Sewell J.C., Scott A.G., Orchard I.,
RA   Rehfeld J.F., Thorpe A.;
RT   "Isolation, structure, and activity of -Phe-Met-Arg-Phe-NH2
RT   neuropeptides (designated calliFMRFamides) from the blowfly Calliphora
RT   vomitoria.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:2326-2330(1992).
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- SIMILARITY: Belongs to the FARP (FMRFamide related peptide)
CC       family.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
```

```
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------------
DR   PIR; B44787; B44787.
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7       7       Phenylalanine amide.
SQ   SEQUENCE   7 AA;  926 MW;  69D40699C44AB700 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          6 R 6


RESULT 39
HCYB_CONCC
ID   HCYB_CONCC     STANDARD;      PRT;     7 AA.
AC   P84620;
DT   13-SEP-2005 (Rel. 48, Created)
DT   13-SEP-2005 (Rel. 48, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Hemocyanin subunit B (CCH-B) (Fragment).
OS   Concholepas concholepas (Barnacle rock-shell).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Orthogastropoda;
OC   Apogastropoda; Caenogastropoda; Sorbeoconcha; Hypsogastropoda;
OC   Neogastropoda; Muricoidea; Muricidae; Concholepas.
OX   NCBI_TaxID=137544;
RN   [1]
RP   PROTEIN SEQUENCE, COFACTOR, SUBUNIT, SUBCELLULAR LOCATION, AND TISSUE
RP   SPECIFICITY.
RC   TISSUE=Hemolymph;
RX   PubMed=15075320; DOI=10.1074/jbc.M400903200;
RA   De Ioannes P., Moltedo B., Oliva H., Pacheco R., Faunes F.,
RA   De Ioannes A.E., Becker M.I.;
RT   "Hemocyanin of the molluscan Concholepas concholepas exhibits an
RT   unusual heterodecameric array of subunits.";
RL   J. Biol. Chem. 279:26134-26142(2004).
CC   -!- FUNCTION: Hemocyanins are copper-containing oxygen carriers
CC       occurring freely dissolved in the hemolymph of many mollusks and
CC       arthropods.
CC   -!- COFACTOR: Binds 2 copper ions per functional unit.
CC   -!- SUBUNIT: Heterodidecamer composed of A and B subunits, each
CC       containing 8 globular oxygen-binding functional units.
CC       Heterogenous decameric or multidecameric structures may also be
CC       formed.
CC   -!- SUBCELLULAR LOCATION: Extracellular.
CC   -!- TISSUE SPECIFICITY: Hemolymph.
CC   -!- PTM: Forms a thioether bond between 2 amino acids (By similarity).
CC   -!- SIMILARITY: Belongs to the tyrosinase family. Hemocyanin
CC       subfamily.
CC   ---------------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------------
DR   PROSITE; PS00209; HEMOCYANIN_1; PARTIAL.
DR   PROSITE; PS00210; HEMOCYANIN_2; PARTIAL.
DR   PROSITE; PS00497; TYROSINASE_1; PARTIAL.
DR   PROSITE; PS00498; TYROSINASE_2; PARTIAL.
KW   Copper; Direct protein sequencing; Hemolymph; Metal-binding;
KW   Oxygen transport; Thioether bond; Transport.
FT   NON_TER       7       7
SQ   SEQUENCE   7 AA;  855 MW;  6AB2C443241AE740 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          3 R 3


RESULT 40
IPYR_CANAL
ID   IPYR_CANAL     STANDARD;      PRT;     7 AA.
AC   P83777;
DT   01-FEB-2005 (Rel. 46, Created)
DT   01-FEB-2005 (Rel. 46, Last sequence update)
DT   01-FEB-2005 (Rel. 46, Last annotation update)
DE   Inorganic pyrophosphatase (EC 3.6.1.1) (Pyrophosphate phospho-
DE   hydrolase) (PPase) (Fragment).
GN   Name=IPP1;
OS   Candida albicans (Yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5476;
RN   [1]
RP   PROTEIN SEQUENCE, SUBCELLULAR LOCATION, AND ANTIGENICITY.
RC   STRAIN=SC5314; TISSUE=Protoplast;
RX   PubMed=15378761; DOI=10.1002/pmic.200400903;
RA   Pitarch A., Abian J., Carrascal M., Sanchez M., Nombela C., Gil C.;
RT   "Proteomics-based identification of novel Candida albicans antigens
RT   for diagnosis of systemic candidiasis in patients with underlying
RT   hematological malignancies.";
RL   Proteomics 4:3084-3106(2004).
CC   -!- CATALYTIC ACTIVITY: Diphosphate + H(2)O = 2 phosphate.
CC   -!- COFACTOR: Binds 4 magnesium ions per subunit. Other metal ions can
CC       support activity, but at a lower rate. Two magnesium ions are
CC       required for the activation of the enzyme and are present before
CC       substrate binds, two additional magnesium ions form complexes with
CC       substrate and product (By similarity).
CC   -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC   -!- MISCELLANEOUS: Has antigenic properties. Elicits a specific immune
CC       response in systemic candidiasis human patients undergoing
CC       malignant hematological disorders.
CC   -!- SIMILARITY: Belongs to the PPase family.
CC   ---------------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------------
DR   InterPro; IPR008162; Pyrophosphatase.
DR   PROSITE; PS00387; PPASE; PARTIAL.
KW   Antigen; Direct protein sequencing; Hydrolase; Magnesium;
KW   Metal-binding.
FT   NON_TER       1       1
FT   NON_TER       7       7
SQ   SEQUENCE   7 AA;  923 MW;  7409D37B1451ADB0 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          7 R 7


RESULT 41
LANC_CARUI
ID   LANC_CARUI     STANDARD;      PRT;     7 AA.
AC   P36960;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   13-SEP-2005 (Rel. 48, Last annotation update)
DE   Lantibiotic carnocin UI49 (Fragment).
```

```
OS   Carnobacterium sp. (strain UI49).
OC   Bacteria; Firmicutes; Lactobacillales; Carnobacteriaceae;
OC   Carnobacterium.
OX   NCBI_TaxID=35782;
RN   [1]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=92321768; PubMed=1622206;
RA   Stoffels G., Nissen-Meyer J., Gudmundsdottir A., Sletten K., Holo H.,
RA   Nes I.F.;
RT   "Purification and characterization of a new bacteriocin isolated from
RT   a Carnobacterium sp.";
RL   Appl. Environ. Microbiol. 58:1417-1422(1992).
CC   -!- FUNCTION: Lanthionine-containing peptide antibiotic (lantibiotic).
CC       Active on Gram-positive bacteria.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Antibiotic; Antimicrobial; Bacteriocin; Direct protein sequencing;
KW   Lantibiotic.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  786 MW;  741776D05B05B810 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          7 R 7


RESULT 42
TY51_LITRU
ID   TY51_LITRU     STANDARD;      PRT;     7 AA.
AC   P82065;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   01-FEB-2005 (Rel. 46, Last annotation update)
DE   Tryptophyllin-5.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Hyloidea; Hylidae;
OC   Pelodryadinae; Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   PROTEIN SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=Skin secretion;
RA   Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australin red tree frog
RT   'Litoria rubella'. The skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: May act as a neuromodulator or neurotransmitter.
CC   -!- SUBCELLULAR LOCATION: Secreted.
CC   -!- TISSUE SPECIFICITY: Expressed by the skin dorsal glands.
CC   -!- MASS SPECTROMETRY: MW=965; METHOD=FAB; RANGE=1-7; NOTE=Ref.1.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Amidation; Amphibian defense peptide; Direct protein sequencing;
KW   Neuropeptide; Pyrrolidone carboxylic acid.
FT   MOD_RES       1      1       Pyrrolidone carboxylic acid.
FT   MOD_RES       7      7       Arginine amide.
SQ   SEQUENCE   7 AA;  983 MW;  7401E9D3676046B0 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          7 R 7


RESULT 43
UF04_MOUSE
ID   UF04_MOUSE     STANDARD;      PRT;     7 AA.
AC   P38642;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Unknown protein from 2D-PAGE of fibroblasts (P46) (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   TISSUE=Fibroblast;
RX   MEDLINE=95009907; PubMed=7523108;
RA   Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;
RT   "Separation and sequencing of familiar and novel murine proteins using
RT   preparative two-dimensional gel electrophoresis.";
RL   Electrophoresis 15:735-745(1994).
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
CC       protein is: 5.0, its MW is: 46 kDa.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Direct protein sequencing.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  766 MW;  68640AB777632700 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          6 R 6


RESULT 44
UH11_RAT
ID   UH11_RAT       STANDARD;      PRT;     7 AA.
AC   P56576;
DT   15-DEC-1998 (Rel. 37, Created)
DT   15-DEC-1998 (Rel. 37, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Unknown protein from 2D-PAGE of heart tissue (Spot P11) (Fragment).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC   Muroidea; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   STRAIN=Wistar; TISSUE=Heart;
RA   Li X.-P., Pleissner K.-P., Scheler C., Regitz-Zagrosek V., Salikov J.,
RA   Jungblut P.R.;
RL   Submitted (SEP-1998) to Swiss-Prot.
CC   -!- MISCELLANEOUS: On the 2D-gel the determined pI of this unknown
```

```
CC       protein is: 8.5, its MW is: 42 kDa.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
KW   Direct protein sequencing.
FT   UNSURE        2      2       S or A.
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  775 MW;  6866DB040DC5A6B0 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          4 R 4


RESULT 45
WWA1_ACHFU
ID   WWA1_ACHFU     STANDARD;      PRT;     7 AA.
AC   P35919;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   WWamide-1.
OS   Achatina fulica (Giant African snail).
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC   Sigmurethra; Achatinoidea; Achatinidae; Achatina.
OX   NCBI_TaxID=6530;
RN   [1]
RP   PROTEIN SEQUENCE.
RC   TISSUE=Ganglion;
RX   MEDLINE=93265912; PubMed=8495720; DOI=10.1016/0014-5793(93)81458-C;
RA   Minakata H., Ikeda T., Muneoka Y., Kobayashi M., Nomoto K.;
RT   "WWamide-1, -2 and -3: novel neuromodulatory peptides isolated from
RT   ganglia of the African giant snail, Achatina fulica.";
RL   FEBS Lett. 323:104-108(1993).
CC   -!- FUNCTION: Exhibits modulatory effects on the peripheral nervous
CC       system. Inhibits activity on a central neuron.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   PIR; S33245; S33245.
KW   Amidation; Direct protein sequencing; Neuropeptide.
FT   MOD_RES       7      7       Tryptophan amide.
SQ   SEQUENCE   7 AA;  993 MW;  7362D5B69B041310 CRC64;

  Query Match             55.6%;  Score 5;  DB 1;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          2 R 2


RESULT 46
Q95945_YEAST
ID   Q95945_YEAST PRELIMINARY;      PRT;     7 AA.
AC   Q95945;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Inside intron 5 (Fragment).
OS   Saccharomyces cerevisiae (Baker's yeast).
OG   Mitochondrion.
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=D273-10B;
RX   MEDLINE=81069885; PubMed=6254986;
RA   Bonitz S.G., Coruzzi G., Thalenfeld B.E., Tzagoloff A., Macino G.;
RT   "Assembly of the mitochondrial membrane system. Structure and
RT   nucleotide sequence of the gene coding for subunit 1 of yeast
RT   cytochrome oxidase.";
RL   J. Biol. Chem. 255:11927-11941(1980).
DR   EMBL; V00694; CAA24066.1; -; Genomic_DNA.
DR   GO; GO:0005739; C:mitochondrion; IEA.
KW   Mitochondrion.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  859 MW;  75B7232362CDC460 CRC64;

  Query Match             55.6%;  Score 5;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          1 R 1


RESULT 47
Q15903_HUMAN
ID   Q15903_HUMAN PRELIMINARY;      PRT;     7 AA.
AC   Q15903;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   Homo sapiens (clone XP7E7B) (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Placenta;
RA   Lee C.-C., Yazdani A., Wehnert M., Bailey J., Couch L., Xiong M.,
RA   Coolbaugh M.I., Chinault C.A., Baldini A., Lindsay E.A., Zhao Z.-Y.,
RA   Caskey C.T.H.;
RT   "Isolation of chromosome-specific genes by reciprocal probing of
RT   arrayed cDNAs and cosmid libraries.";
RL   Hum. Mol. Genet. 0:0-0(1995).
DR   EMBL; L32082; AAA73893.1; -; mRNA.
FT   NON_TER       1      1
FT   NON_TER       7      7
SQ   SEQUENCE   7 AA;  849 MW;  6B040339CDD33DB0 CRC64;

  Query Match             55.6%;  Score 5;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
               |
Db          6 R 6


RESULT 48
Q8NHH7_HUMAN
ID   Q8NHH7_HUMAN PRELIMINARY;      PRT;     7 AA.
AC   Q8NHH7;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Mini-cistron.
```

```
GN   Name=NHE3;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Malakooti J., Ramaswamy K.;
RL   Submitted (JUN-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF282824; AAM53436.1; -; Genomic_DNA.
SQ   SEQUENCE   7 AA;  842 MW;  74072DC772D406F0 CRC64;

  Query Match             55.6%;  Score 5;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 49
Q8TAQ4_HUMAN
ID   Q8TAQ4_HUMAN PRELIMINARY;      PRT;     7 AA.
AC   Q8TAQ4;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   YAP1 protein.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Uterus;
RX   MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Uterus;
RA   Strausberg R.;
RL   Submitted (APR-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC026212; AAH26212.2; -; mRNA.
SQ   SEQUENCE   7 AA;  848 MW;  6AB2D1B6C2D406F0 CRC64;

  Query Match             55.6%;  Score 5;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          2 R 2


RESULT 50
O98866_SPIOL
ID   O98866_SPIOL PRELIMINARY;      PRT;     7 AA.
AC   O98866;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Cytochrome b/f subunit IV (Fragment).
OS   Spinacia oleracea (Spinach).
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC   Caryophyllales; Amaranthaceae; Spinacia.
OX   NCBI_TaxID=3562;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=86120353; PubMed=3003688;
RA   Sijben-Mueller G., Hallick R.B., Alt J., Westhoff P., Herrmann R.G.;
RT   "Spinach plastid genes coding for initiation factor IF-1, ribosomal
RT   protein S11 and RNA polymerase alpha-subunit.";
RL   Nucleic Acids Res. 14:1029-1044(1986).
DR   EMBL; X03496; CAA27215.1; -; Genomic_DNA.
DR   GO; GO:0009507; C:chloroplast; IEA.
KW   Chloroplast.
FT   NON_TER       1      1
SQ   SEQUENCE   7 AA;  907 MW;  644729D77409C420 CRC64;

  Query Match             55.6%;  Score 5;  DB 2;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 0;
  Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 R 1
              |
Db          3 R 3


Search completed: January 25, 2006, 18:41:14
Job time : 92 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:32:42 ; Search time 78.5 Seconds
(without alignments)
27.986 Million cell updates/sec

Title: US-10-771-242-295
Perfect score: 21
Sequence: 1 RRLNX 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	95.2	5	9	ADZ71770	Adz71770 p21-deriv
2	20	95.2	5	9	ADZ72025	Adz72025 p21-deriv
3	20	95.2	5	9	ADZ71763	Adz71763 p21-deriv
4	20	95.2	5	9	ADZ72027	Adz72027 p21-deriv
5	20	95.2	5	9	ADZ72107	Adz72107 p21-deriv
6	20	95.2	5	9	ADZ71617	Adz71617 p21-deriv
7	20	95.2	5	9	ADZ71663	Adz71663 p21-deriv
8	20	95.2	5	9	ADZ71764	Adz71764 p21-deriv
9	20	95.2	5	9	ADZ71765	Adz71765 p21-deriv
10	20	95.2	5	9	ADZ71746	Adz71746 p21-deriv
11	20	95.2	5	9	ADZ71618	Adz71618 p21-deriv
12	20	95.2	5	9	ADZ71662	Adz71662 p21-deriv
13	20	95.2	5	9	ADZ71664	Adz71664 p21-deriv
14	20	95.2	5	9	ADZ72026	Adz72026 p21-deriv
15	20	95.2	5	9	ADZ72106	Adz72106 p21-deriv
16	20	95.2	5	9	ADZ71747	Adz71747 p21-deriv
17	20	95.2	5	9	ADZ71748	Adz71748 p21-deriv
18	20	95.2	6	2	AAW56904	Aaw56904 Enzyme in
19	20	95.2	7	4	AAM46084	Aam46084 H11 bindi
20	20	95.2	7	4	AAM46173	Aam46173 H11 bindi
21	20	95.2	7	4	AAM44565	Aam44565 H11 bindi
22	20	95.2	7	4	AAM44575	Aam44575 H11 bindi
23	20	95.2	7	4	AAM43976	Aam43976 H11 bindi
24	20	95.2	7	4	AAM44570	Aam44570 H11 bindi
25	20	95.2	7	4	AAM46480	Aam46480 H11 bindi
26	20	95.2	7	4	AAM44560	Aam44560 H11 bindi
27	20	95.2	7	4	AAM45974	Aam45974 H11 bindi
28	20	95.2	7	4	AAM46528	Aam46528 H11 bindi
29	20	95.2	7	4	AAM46485	Aam46485 H11 bindi
30	20	95.2	7	4	AAM45665	Aam45665 H11 bindi
31	20	95.2	7	4	AAM46523	Aam46523 H11 bindi
32	20	95.2	7	4	AAM46526	Aam46526 H11 bindi
33	20	95.2	8	2	AAW57007	Aaw57007 Enzyme in
34	20	95.2	8	4	AAU05744	Aau05744 p21 C-ter
35	20	95.2	8	9	ADZ71972	Adz71972 p21-deriv
36	20	95.2	8	9	ADZ71569	Adz71569 p21-deriv
37	20	95.2	9	5	AAE18742	Aae18742 Human leu
38	20	95.2	9	5	AAE18741	Aae18741 Human leu
39	20	95.2	9	5	AAE18780	Aae18780 Human leu
40	20	95.2	11	4	AAB80739	Aab80739 Semenogel
41	20	95.2	11	7	ABW01200	Abw01200 Saccharom
42	20	95.2	11	8	ADP44126	Adp44126 Yeast tra
43	20	95.2	12	4	AAB59999	Aab59999 Internali
44	20	95.2	12	7	ADA88834	Ada88834 Internali
45	20	95.2	12	8	ADQ15921	Adq15921 Human Mas
46	20	95.2	14	2	AAR47006	Aar47006 LAR prote
47	20	95.2	15	2	AAW42267	Aaw42267 Biotinyla
48	20	95.2	15	2	AAW15509	Aaw15509 Peptide M
49	20	95.2	15	4	ABB76940	Abb76940 Rat VG51-
50	20	95.2	15	5	AAE18774	Aae18774 Human leu
51	20	95.2	19	5	AAU93571	Aau93571 Granulocy
52	20	95.2	20	6	ADA08204	Ada08204 Human CRK
53	20	95.2	20	7	ADH47647	Adh47647 Human lun
54	20	95.2	20	7	ADH47648	Adh47648 Human lun
55	20	95.2	20	8	ADJ21567	Adj21567 Human lun
56	20	95.2	20	8	ADJ21566	Adj21566 Human lun
57	20	95.2	24	2	AAR85034	Aar85034 Peptide r
58	20	95.2	28	2	AAY20367	Aay20367 Human mic
59	20	95.2	32	2	AAR85789	Aar85789 Peptide r
60	20	95.2	33	2	AAY48432	Aay48432 Human pro
61	20	95.2	33	7	ABM79022	Abm79022 BK channe
62	20	95.2	34	3	AAB34526	Aab34526 Human sec
63	20	95.2	35	4	ABB40086	Abb40086 Peptide #
64	20	95.2	35	4	AAM33720	Aam33720 Peptide #
65	20	95.2	35	4	AAM73524	Aam73524 Human bon
66	20	95.2	35	4	AAM60842	Aam60842 Human bra
67	20	95.2	35	4	ABG55251	Abg55251 Human liv
68	20	95.2	35	5	ABG43385	Abg43385 Human pep
69	20	95.2	36	2	AAY59938	Aay59938 Human myo
70	20	95.2	37	2	AAR85022	Aar85022 Peptide r
71	20	95.2	40	7	ABR82177	Abr82177 Source 27
72	20	95.2	40	7	ABR82163	Abr82163 Source 13
73	20	95.2	41	4	ABB03590	Abb03590 Human mus
74	20	95.2	41	6	ABU12884	Abu12884 Novel hum
75	20	95.2	41	8	ADJ28910	Adj28910 Human mus
76	20	95.2	47	3	AAG45682	Aag45682 Arabidops
77	20	95.2	47	4	ABB17914	Abb17914 Human ner
78	20	95.2	47	9	ADY66570	Ady66570 S. manson
79	20	95.2	48	8	ADS07330	Ads07330 Staphyloc
80	20	95.2	49	4	AAM23505	Aam23505 Human EST
81	20	95.2	49	9	ADY30821	Ady30821 Thale cre
82	20	95.2	50	4	ABB42086	Abb42086 Peptide #
83	20	95.2	50	4	AAM35888	Aam35888 Peptide #
84	20	95.2	50	4	AAM75778	Aam75778 Human bon
85	20	95.2	50	4	AAU46056	Aau46056 Propionib
86	20	95.2	50	4	AAM62966	Aam62966 Human bra
87	20	95.2	50	4	ABG57516	Abg57516 Human liv
88	20	95.2	50	5	ABG45252	Abg45252 Human pep
89	20	95.2	50	6	ABM42575	Abm42575 Propionib
90	20	95.2	51	4	ABB32056	Abb32056 Peptide #
91	20	95.2	51	4	ABB22594	Abb22594 Protein #
92	20	95.2	51	4	AAU60742	Aau60742 Propionib
93	20	95.2	51	4	ABG52125	Abg52125 Human liv
94	20	95.2	51	5	ABP02714	Abp02714 Human ORF
95	20	95.2	51	5	ABG40076	Abg40076 Human pep
96	20	95.2	51	6	ABM57261	Abm57261 Propionib
97	20	95.2	52	8	ADS07331	Ads07331 Staphyloc

98	20	95.2	53	4	AAM85803	Aam85803	Human imm
99	20	95.2	53	4	AAU47841	Aau47841	Propionib
100	20	95.2	53	6	ABM44360	Abm44360	Propionib
101	20	95.2	53	8	ADJ66980	Adj66980	Human sec
102	20	95.2	54	3	AAG03645	Aag03645	Human sec
103	20	95.2	54	4	AAU46799	Aau46799	Propionib
104	20	95.2	54	4	AAU47820	Aau47820	Propionib
105	20	95.2	54	6	ABM44339	Abm44339	Propionib
106	20	95.2	54	6	ABM43318	Abm43318	Propionib
107	20	95.2	55	4	AAU60505	Aau60505	Propionib
108	20	95.2	55	6	ABM57024	Abm57024	Propionib
109	20	95.2	56	4	AAM20689	Aam20689	Peptide #
110	20	95.2	56	4	ABB42078	Abb42078	Peptide #
111	20	95.2	56	4	AAM35880	Aam35880	Peptide #
112	20	95.2	56	4	ABB25675	Abb25675	Protein #
113	20	95.2	56	4	AAM75770	Aam75770	Human bon
114	20	95.2	56	4	AAU59928	Aau59928	Propionib
115	20	95.2	56	4	AAM62958	Aam62958	Human bra
116	20	95.2	56	4	ABG57510	Abg57510	Human liv
117	20	95.2	56	5	ABG45247	Abg45247	Human pep
118	20	95.2	56	6	ABM56447	Abm56447	Propionib
119	20	95.2	58	4	AAO04670	Aao04670	Human pol
120	20	95.2	58	5	ABP07646	Abp07646	Human ORF
121	20	95.2	58	7	ADG10562	Adg10562	Human STA
122	20	95.2	59	3	AAG01585	Aag01585	Human sec
123	20	95.2	59	4	AAU49510	Aau49510	Propionib
124	20	95.2	59	6	ABM46029	Abm46029	Propionib
125	20	95.2	60	5	ADK34339	Adk34339	Novel hum
126	20	95.2	61	3	AAG58739	Aag58739	Arabidops
127	20	95.2	61	3	AAG37096	Aag37096	Arabidops
128	20	95.2	61	4	AAU65071	Aau65071	Propionib
129	20	95.2	61	4	AAU50814	Aau50814	Propionib
130	20	95.2	61	6	ABM47333	Abm47333	Propionib
131	20	95.2	61	6	ABM61590	Abm61590	Propionib
132	20	95.2	61	6	ABP76329	Abp76329	Human GEN
133	20	95.2	61	8	ADR96519	Adr96519	Novel S.
134	20	95.2	61	9	AEA60389	Aea60389	Streptoco
135	20	95.2	62	4	AAG74875	Aag74875	Human col
136	20	95.2	62	4	AAU59323	Aau59323	Propionib
137	20	95.2	62	4	AAU59442	Aau59442	Propionib
138	20	95.2	62	5	ABP02008	Abp02008	Human ORF
139	20	95.2	62	6	ABM55961	Abm55961	Propionib
140	20	95.2	62	6	ABM55842	Abm55842	Propionib
141	20	95.2	63	4	AAU59231	Aau59231	Propionib
142	20	95.2	63	6	ABM55750	Abm55750	Propionib
143	20	95.2	63	6	ABU22642	Abu22642	Protein e
144	20	95.2	63	6	ABU21558	Abu21558	Protein e
145	20	95.2	64	4	AAU48935	Aau48935	Propionib
146	20	95.2	64	6	ABM45454	Abm45454	Propionib
147	20	95.2	64	8	ABO57564	Abo57564	Human gen
148	20	95.2	65	3	AAB32396	Aab32396	Human sec
149	20	95.2	65	4	AAU54243	Aau54243	Propionib
150	20	95.2	65	4	ABB17150	Abb17150	Human ner
151	20	95.2	65	6	ABM50762	Abm50762	Propionib
152	20	95.2	65	6	ABB99052	Abb99052	S. coelic
153	20	95.2	65	7	ADH87393	Adh87393	Enterococ
154	20	95.2	66	4	AAM94716	Aam94716	Human rep
155	20	95.2	66	4	AAO01640	Aao01640	Human pol
156	20	95.2	66	4	AAU52583	Aau52583	Propionib
157	20	95.2	66	4	AAU22743	Aau22743	Human pro
158	20	95.2	66	6	ABM49102	Abm49102	Propionib
159	20	95.2	66	7	ADJ09316	Adj09316	Human pro
160	20	95.2	67	4	AAU21127	Aau21127	Human nov
161	20	95.2	68	4	AAM92836	Aam92836	Human dig
162	20	95.2	68	4	AAM90490	Aam90490	Human imm
163	20	95.2	68	4	ABB11674	Abb11674	Human rib
164	20	95.2	68	4	AAM80273	Aam80273	Human pro
165	20	95.2	68	7	ADC33022	Adc33022	Human nov
166	20	95.2	69	5	ABP34005	Abp34005	Human ORF
167	20	95.2	70	4	AAU51980	Aau51980	Propionib
168	20	95.2	70	6	ABM48499	Abm48499	Propionib
169	20	95.2	71	2	AAY76577	Aay76577	Human ova
170	20	95.2	71	4	AAU54982	Aau54982	Propionib
171	20	95.2	71	6	ABM51501	Abm51501	Propionib
172	20	95.2	72	5	ABP31506	Abp31506	Human ORF
173	20	95.2	72	8	ADL04327	Adl04327	M. catarr
174	20	95.2	73	3	AAB14406	Aab14406	Human lip
175	20	95.2	73	4	ABB11006	Abb11006	Human rev
176	20	95.2	73	4	AAE03867	Aae03867	Human gen
177	20	95.2	73	5	ABG64571	Abg64571	Human alb
178	20	95.2	73	7	ADC00887	Adc00887	Enterohae
179	20	95.2	73	7	ADC00210	Adc00210	Enterohae
180	20	95.2	73	7	ADC00577	Adc00577	Enterohae
181	20	95.2	73	8	ADL77838	Adl77838	Albumin f
182	20	95.2	74	3	AAY74684	Aay74684	Neisseria
183	20	95.2	74	6	ABU70734	Abu70734	Human adi
184	20	95.2	75	5	ABP29218	Abp29218	Streptoco
185	20	95.2	75	6	ADB10678	Adb10678	Alloiococ
186	20	95.2	76	3	AAB14404	Aab14404	Rat RhoGA
187	20	95.2	77	4	AAM38481	Aam38481	Peptide #
188	20	95.2	77	4	AAM65585	Aam65585	Human bra
189	20	95.2	77	5	ABP00402	Abp00402	Human ORF
190	20	95.2	77	8	ABO55675	Abo55675	Human gen
191	20	95.2	77	8	ADN99383	Adn99383	Novel hum
192	20	95.2	78	4	ABG29189	Abg29189	Novel hum
193	20	95.2	78	5	ABP00882	Abp00882	Human ORF
194	20	95.2	79	5	ADK34737	Adk34737	Novel hum
195	20	95.2	80	3	AAG58738	Aag58738	Arabidops
196	20	95.2	80	5	ABP03115	Abp03115	Human ORF
197	20	95.2	80	7	ADC00984	Adc00984	Enterohae
198	20	95.2	80	7	ADH88312	Adh88312	Enterococ
199	20	95.2	80	8	ABO57578	Abo57578	Human gen
200	20	95.2	81	3	AAY91643	Aay91643	Human sec
201	20	95.2	81	4	AAM84170	Aam84170	Human imm
202	20	95.2	81	4	AAU59644	Aau59644	Propionib
203	20	95.2	81	6	ABM56163	Abm56163	Propionib
204	20	95.2	81	8	ADL71721	Adl71721	Novel hum
205	20	95.2	81	8	ADT57031	Adt57031	Plant pol
206	20	95.2	82	3	AAY91557	Aay91557	Human sec
207	20	95.2	82	4	AAU21252	Aau21252	Human nov
208	20	95.2	82	8	ADL71630	Adl71630	Novel hum
209	20	95.2	83	4	AAO01170	Aao01170	Human pol
210	20	95.2	83	4	AAU39795	Aau39795	Propionib
211	20	95.2	83	4	AAU61154	Aau61154	Propionib
212	20	95.2	83	6	ABM57673	Abm57673	Propionib
213	20	95.2	83	6	ABM36314	Abm36314	Propionib
214	20	95.2	84	3	AAG20426	Aag20426	Arabidops
215	20	95.2	85	3	AAB32591	Aab32591	Eucalyptu
216	20	95.2	85	5	ABP00233	Abp00233	Human ORF
217	20	95.2	85	5	AAE23312	Aae23312	Human p54
218	20	95.2	85	5	AAE23318	Aae23318	Human p85
219	20	95.2	85	7	ADD36316	Add36316	D melanog
220	20	95.2	85	7	ADD36330	Add36330	D melanog
221	20	95.2	85	8	ADQ08953	Adq08953	D. melano
222	20	95.2	85	8	ADQ08939	Adq08939	D. melano
223	20	95.2	85	9	ADW46118	Adw46118	Fruit fly
224	20	95.2	85	9	ADW46132	Adw46132	Fruit fly
225	20	95.2	86	4	AAO10340	Aao10340	Human pol
226	20	95.2	86	4	AAU39146	Aau39146	Propionib
227	20	95.2	86	4	AAU40456	Aau40456	Propionib
228	20	95.2	86	6	ABM35665	Abm35665	Propionib
229	20	95.2	86	6	ABM36975	Abm36975	Propionib
230	20	95.2	86	6	ABU70552	Abu70552	Human adi
231	20	95.2	86	9	ADZ10664	Adz10664	N. europe
232	20	95.2	87	2	AAW35568	Aaw35568	Borna dis
233	20	95.2	87	3	AAY79287	Aay79287	Borna dis
234	20	95.2	87	5	ABB78170	Abb78170	Amino aci
235	20	95.2	87	7	ADD36276	Add36276	D melanog
236	20	95.2	87	8	ADL05617	Adl05617	M. catarr
237	20	95.2	87	8	ADQ08899	Adq08899	Anopheles
238	20	95.2	87	9	ADW46078	Adw46078	Fruit fly
239	20	95.2	88	5	ABP28776	Abp28776	Streptoco
240	20	95.2	88	8	ADV89222	Adv89222	Streptoco
241	20	95.2	88	8	ADV80475	Adv80475	Streptoco
242	20	95.2	88	8	ADV82536	Adv82536	Streptoco
243	20	95.2	89	4	AAU45557	Aau45557	Propionib

244	20	95.2	89	6	ABM42076	Abm42076	Propionib	317	20	95.2	111	5	ABP41090	Abp41090	Human ova
245	20	95.2	90	3	AAB13265	Aab13265	Caenorhab	318	20	95.2	111	7	ABM86150	Abm86150	Rice abio
246	20	95.2	90	4	AAB78982	Aab78982	C. glutam	319	20	95.2	111	8	ADL81889	Adl81889	P. aerugi
247	20	95.2	90	5	ABB78168	Abb78168	Amino aci	320	20	95.2	112	4	AAE01912	Aae01912	Arabidops
248	20	95.2	90	8	ADX78594	Adx78594	Plant ful	321	20	95.2	112	4	AAE02467	Aae02467	Arabidops
249	20	95.2	91	4	AAG75213	Aag75213	Human col	322	20	95.2	112	7	ADD55784	Add55784	Thalecres
250	20	95.2	91	6	ABU70576	Abu70576	Human adi	323	20	95.2	112	7	ADD30554	Add30554	Plant yie
251	20	95.2	92	4	AAM25520	Aam25520	Human pro	324	20	95.2	112	7	ADE37115	Ade37115	Plant yie
252	20	95.2	93	7	ADC94062	Adc94062	E. faeciu	325	20	95.2	112	8	ADI41575	Adi41575	Plant tra
253	20	95.2	94	4	AAG74198	Aag74198	Human col	326	20	95.2	112	8	ADJ56850	Adj56850	Arabidops
254	20	95.2	94	4	ABB70492	Abb70492	Drosophil	327	20	95.2	112	8	ADO03337	Ado03337	Thalecres
255	20	95.2	94	4	AAU60007	Aau60007	Propionib	328	20	95.2	112	8	ADO03603	Ado03603	Thalecres
256	20	95.2	94	4	ABG11744	Abg11744	Novel hum	329	20	95.2	112	8	ADO01729	Ado01729	Thalecres
257	20	95.2	94	5	ABP25808	Abp25808	Streptoco	330	20	95.2	112	8	ADO62911	Ado62911	Transcrip
258	20	95.2	94	6	ABM56526	Abm56526	Propionib	331	20	95.2	112	8	ADS16859	Ads16859	Thale cre
259	20	95.2	94	8	ADV89513	Adv89513	Streptoco	332	20	95.2	112	9	AEA26445	Aea26445	Stress to
260	20	95.2	94	8	ADV82929	Adv82929	Streptoco	333	20	95.2	113	4	AAO09987	Aao09987	Human pol
261	20	95.2	94	8	ADV80766	Adv80766	Streptoco	334	20	95.2	113	6	ADB10752	Adb10752	Alloiococ
262	20	95.2	95	3	AAY79286	Aay79286	Borna dis	335	20	95.2	113	7	ADM25417	Adm25417	Hyperther
263	20	95.2	95	3	AAG51932	Aag51932	Arabidops	336	20	95.2	114	4	AAG91809	Aag91809	C glutami
264	20	95.2	95	3	AAG20632	Aag20632	Arabidops	337	20	95.2	115	3	AAY91642	Aay91642	Human sec
265	20	95.2	95	4	ABB11442	Abb11442	Human sec	338	20	95.2	115	4	AAG91827	Aag91827	C glutami
266	20	95.2	95	4	AAU49110	Aau49110	Propionib	339	20	95.2	115	5	ABP08923	Abp08923	Human ORF
267	20	95.2	95	4	ABG18078	Abg18078	Novel hum	340	20	95.2	115	6	ABU70451	Abu70451	Human adi
268	20	95.2	95	6	ABM45629	Abm45629	Propionib	341	20	95.2	115	8	ADL71720	Adl71720	Novel hum
269	20	95.2	95	8	ADX93779	Adx93779	Plant ful	342	20	95.2	115	8	ADX70698	Adx70698	Plant ful
270	20	95.2	95	8	ADY24463	Ady24463	Plant ful	343	20	95.2	116	5	ABP32944	Abp32944	Human pho
271	20	95.2	96	3	AAB29476	Aab29476	Burkholde	344	20	95.2	117	4	AAM82867	Aam82867	Human imm
272	20	95.2	96	4	ABG21787	Abg21787	Novel hum	345	20	95.2	117	4	AAB95436	Aab95436	Human pro
273	20	95.2	97	2	AAY37141	Aay37141	Amino aci	346	20	95.2	118	5	ABP27226	Abp27226	Streptoco
274	20	95.2	97	4	AAU55876	Aau55876	Propionib	347	20	95.2	118	6	ABU46798	Abu46798	Protein e
275	20	95.2	97	6	ABM52395	Abm52395	Propionib	348	20	95.2	118	6	ABP76051	Abp76051	Human GEN
276	20	95.2	97	7	ABO79693	Abo79693	Pseudomon	349	20	95.2	118	8	ADX75855	Adx75855	Plant ful
277	20	95.2	98	3	AAG34149	Aag34149	Zea mays	350	20	95.2	119	3	AAY75271	Aay75271	Neisseria
278	20	95.2	98	4	ABB63383	Abb63383	Drosophil	351	20	95.2	119	5	ABP40144	Abp40144	Staphyloc
279	20	95.2	98	4	AAU47172	Aau47172	Propionib	352	20	95.2	119	7	ABO77643	Abo77643	Pseudomon
280	20	95.2	98	5	ABP07670	Abp07670	Human ORF	353	20	95.2	119	8	ADS07188	Ads07188	Staphyloc
281	20	95.2	98	6	ABM43691	Abm43691	Propionib	354	20	95.2	120	3	AAG49392	Aag49392	Arabidops
282	20	95.2	99	5	ABR40491	Abr40491	Human sec	355	20	95.2	120	4	AAM85739	Aam85739	Human imm
283	20	95.2	99	5	ABR40417	Abr40417	Human sec	356	20	95.2	120	4	ABB12278	Abb12278	Human sec
284	20	95.2	99	5	ABP31616	Abp31616	Human ORF	357	20	95.2	120	4	ABG11980	Abg11980	Novel hum
285	20	95.2	100	3	AAG01140	Aag01140	Human sec	358	20	95.2	121	2	AAW77543	Aaw77543	Staphyloc
286	20	95.2	100	3	AAG55172	Aag55172	Arabidops	359	20	95.2	121	3	AAG49391	Aag49391	Arabidops
287	20	95.2	100	4	ABB66078	Abb66078	Drosophil	360	20	95.2	121	4	AAU48378	Aau48378	Propionib
288	20	95.2	100	4	AAM91091	Aam91091	Human imm	361	20	95.2	121	4	AAU44395	Aau44395	Propionib
289	20	95.2	100	4	ABG16166	Abg16166	Novel hum	362	20	95.2	121	6	ABM44897	Abm44897	Propionib
290	20	95.2	100	8	ADJ57835	Adj57835	MRB prote	363	20	95.2	121	6	ABM40914	Abm40914	Propionib
291	20	95.2	101	6	ADB11138	Adb11138	Alloiococ	364	20	95.2	122	3	AAG37576	Aag37576	Arabidops
292	20	95.2	102	4	AAU50064	Aau50064	Propionib	365	20	95.2	122	3	AAG49390	Aag49390	Arabidops
293	20	95.2	102	6	ABM46583	Abm46583	Propionib	366	20	95.2	122	5	ABG64896	Abg64896	Human alb
294	20	95.2	104	4	AAO13013	Aao13013	Human pol	367	20	95.2	122	5	ABB77034	Abb77034	Human pro
295	20	95.2	104	4	AAG82569	Aag82569	S. epider	368	20	95.2	122	7	ADC88716	Adc88716	Ribosomal
296	20	95.2	105	4	AAM24351	Aam24351	Human EST	369	20	95.2	122	7	ABO74609	Abo74609	Pseudomon
297	20	95.2	105	7	ADC31355	Adc31355	Human nov	370	20	95.2	122	8	ADL78163	Adl78163	Albumin f
298	20	95.2	106	3	AAG05441	Aag05441	Arabidops	371	20	95.2	123	3	AAG03927	Aag03927	Human sec
299	20	95.2	106	4	ABB66482	Abb66482	Drosophil	372	20	95.2	123	5	ABP64793	Abp64793	Human pro
300	20	95.2	106	4	ABB66481	Abb66481	Drosophil	373	20	95.2	123	7	ADM04917	Adm04917	Human pro
301	20	95.2	106	4	AAU14407	Aau14407	Human nov	374	20	95.2	123	7	ABM87812	Abm87812	Rice abio
302	20	95.2	106	4	AAU14408	Aau14408	Human nov	375	20	95.2	123	8	ABM81082	Abm81082	Tumour-as
303	20	95.2	106	7	ADE09443	Ade09443	Novel pro	376	20	95.2	123	9	ADX08058	Adx08058	Cyclin-de
304	20	95.2	106	8	ADH80726	Adh80726	Human pol	377	20	95.2	124	7	ABM74443	Abm74443	DNA clone
305	20	95.2	106	8	ADH80725	Adh80725	Human pol	378	20	95.2	125	2	AAY85832	Aay85832	S. pneumo
306	20	95.2	107	4	AAU67163	Aau67163	Propionib	379	20	95.2	125	8	ADK46925	Adk46925	Streptoco
307	20	95.2	107	5	ABP68923	Abp68923	Human pol	380	20	95.2	126	3	AAG32992	Aag32992	Arabidops
308	20	95.2	107	5	ABP38042	Abp38042	Staphyloc	381	20	95.2	126	5	ABP33115	Abp33115	Human red
309	20	95.2	107	6	ABM63682	Abm63682	Propionib	382	20	95.2	126	6	ABR40675	Abr40675	Zea mays
310	20	95.2	109	4	AAU37519	Aau37519	Staphyloc	383	20	95.2	126	8	ABM80250	Abm80250	Tumour-as
311	20	95.2	109	8	ADQ66927	Adq66927	Novel hum	384	20	95.2	127	4	AAU64816	Aau64816	Propionib
312	20	95.2	110	4	AAU45128	Aau45128	Propionib	385	20	95.2	127	6	ABM61335	Abm61335	Propionib
313	20	95.2	110	4	AAU32624	Aau32624	Novel hum	386	20	95.2	127	7	ADM26194	Adm26194	Hyperther
314	20	95.2	110	6	ABM41647	Abm41647	Propionib	387	20	95.2	127	8	ADX66219	Adx66219	Plant ful
315	20	95.2	111	3	AAB43863	Aab43863	Human can	388	20	95.2	128	4	AAG98867	Aag98867	E. coli g
316	20	95.2	111	3	AAB56865	Aab56865	Human pro	389	20	95.2	128	5	ADK36004	Adk36004	Novel hum

390	20	95.2	129	5	ABP34766	Abp34766	Human iso
391	20	95.2	129	8	ABM80361	Abm80361	Tumour-as
392	20	95.2	129	9	ABM94297	Abm94297	M. xanthu
393	20	95.2	130	4	ABU53192	Abu53192	Human tes
394	20	95.2	130	8	ADN05496	Adn05496	Antipsori
395	20	95.2	131	3	AAG56232	Aag56232	Arabidops
396	20	95.2	131	3	AAG58827	Aag58827	Arabidops
397	20	95.2	131	4	AAU23240	Aau23240	Novel hum
398	20	95.2	132	8	ADX79464	Adx79464	Plant ful
399	20	95.2	132	8	ADX79474	Adx79474	Plant ful
400	20	95.2	133	4	ABB70513	Abb70513	Drosophil
401	20	95.2	133	7	ABO61598	Abo61598	Klebsiell
402	20	95.2	134	2	AAR32292	Aar32292	Sequence
403	20	95.2	134	3	AAG08210	Aag08210	Arabidops
404	20	95.2	135	3	AAG03345	Aag03345	Human sec
405	20	95.2	135	3	AAG26193	Aag26193	Arabidops
406	20	95.2	135	8	ADX73541	Adx73541	Plant ful
407	20	95.2	136	4	AAU57372	Aau57372	Propionib
408	20	95.2	136	4	AAE04765	Aae04765	Human ves
409	20	95.2	136	6	ABM53891	Abm53891	Propionib
410	20	95.2	136	6	ABU70404	Abu70404	Human adi
411	20	95.2	136	8	ABM80151	Abm80151	Tumour-as
412	20	95.2	137	3	AAG08209	Aag08209	Arabidops
413	20	95.2	137	6	ABU70769	Abu70769	Human adi
414	20	95.2	139	4	ABB43303	Abb43303	Peptide #
415	20	95.2	139	4	AAM37148	Aam37148	Peptide #
416	20	95.2	139	4	AAM77037	Aam77037	Human bon
417	20	95.2	139	4	AAM64211	Aam64211	Human bra
418	20	95.2	139	4	ABG58696	Abg58696	Human liv
419	20	95.2	139	7	ABO76363	Abo76363	Pseudomon
420	20	95.2	139	7	ABO75668	Abo75668	Pseudomon
421	20	95.2	139	9	ADY27665	Ady27665	Human ach
422	20	95.2	140	5	ABG77378	Abg77378	Selected
423	20	95.2	140	7	ADC00833	Adc00833	Enterohae
424	20	95.2	140	8	ADQ16007	Adq16007	hMash-1-d
425	20	95.2	140	9	ADY27672	Ady27672	Human ach
426	20	95.2	141	6	ABU40239	Abu40239	Protein e
427	20	95.2	142	3	AAB59091	Aab59091	Breast an
428	20	95.2	142	4	ABG27260	Abg27260	Novel hum
429	20	95.2	142	7	ABO69587	Abo69587	Pseudomon
430	20	95.2	143	5	AAE17807	Aae17807	Herpes si
431	20	95.2	143	7	ADG74939	Adg74939	Human her
432	20	95.2	144	4	AAU36192	Aau36192	Pseudomon
433	20	95.2	144	5	ABP29247	Abp29247	Streptoco
434	20	95.2	144	6	ABU38249	Abu38249	Protein e
435	20	95.2	144	7	ADC89174	Adc89174	Ribosomal
436	20	95.2	144	7	ADC89161	Adc89161	Ribosomal
437	20	95.2	144	7	ADE60089	Ade60089	Rat Prote
438	20	95.2	144	7	ADE60091	Ade60091	Human Pro
439	20	95.2	144	7	ADD45769	Add45769	Rat Prote
440	20	95.2	144	7	ADE60093	Ade60093	Rat Prote
441	20	95.2	144	7	ADE56508	Ade56508	Human Pro
442	20	95.2	144	7	ADE60095	Ade60095	Human Pro
443	20	95.2	144	7	ADD45771	Add45771	Human Pro
444	20	95.2	145	1	AAP81003	Aap81003	Sequence
445	20	95.2	145	7	ADB85279	Adb85279	Rat DNA b
446	20	95.2	145	7	ADC89175	Adc89175	Ribosomal
447	20	95.2	145	7	ADE56506	Ade56506	Rat Prote
448	20	95.2	145	8	ABM81988	Abm81988	Tumour-as
449	20	95.2	145	8	ADU60158	Adu60158	Housekeep
450	20	95.2	145	8	ADX80197	Adx80197	Plant ful
451	20	95.2	146	4	AAO00377	Aao00377	Human pol
452	20	95.2	146	7	ADE60911	Ade60911	Rat Prote
453	20	95.2	146	7	ADE60915	Ade60915	Rat Prote
454	20	95.2	146	8	ADX90949	Adx90949	Plant ful
455	20	95.2	146	8	ADX90435	Adx90435	Plant ful
456	20	95.2	147	4	ABG25058	Abg25058	Novel hum
457	20	95.2	147	5	ABB97880	Abb97880	Human sec
458	20	95.2	147	5	AAU99297	Aau99297	Pyrococcu
459	20	95.2	147	6	ABU02446	Abu02446	S. pneumo
460	20	95.2	147	8	ADQ15865	Adq15865	Predicted
461	20	95.2	147	8	ADQ15867	Adq15867	YJEE prot
462	20	95.2	149	3	AAB41916	Aab41916	Human ORF
463	20	95.2	149	4	AAU19479	Aau19479	Human dia
464	20	95.2	149	6	ABM69638	Abm69638	Photorhab
465	20	95.2	149	7	ADF59714	Adf59714	Human pol
466	20	95.2	149	7	ABO83188	Abo83188	Pseudomon
467	20	95.2	149	8	ADI67173	Adi67173	Lactobaci
468	20	95.2	149	8	ADR94703	Adr94703	Novel S.
469	20	95.2	149	8	ADU73960	Adu73960	Melampsor
470	20	95.2	149	9	AEA58573	Aea58573	Streptoco
471	20	95.2	150	3	AAG58826	Aag58826	Arabidops
472	20	95.2	150	5	ABJ11280	Abj11280	Yeast sel
473	20	95.2	150	7	ADE60917	Ade60917	Human Pro
474	20	95.2	150	7	ADE60913	Ade60913	Human Pro
475	20	95.2	150	7	ADJ68771	Adj68771	Human hea
476	20	95.2	150	8	ADU73953	Adu73953	Melampsor
477	20	95.2	150	8	ADU73955	Adu73955	Melampsor
478	20	95.2	150	8	ADU73959	Adu73959	Melampsor
479	20	95.2	151	3	AAG23136	Aag23136	Arabidops
480	20	95.2	151	4	AAU59190	Aau59190	Propionib
481	20	95.2	151	4	AAU56629	Aau56629	Propionib
482	20	95.2	151	6	ABP78723	Abp78723	N. gonorr
483	20	95.2	151	6	ABM55709	Abm55709	Propionib
484	20	95.2	151	6	ABM53148	Abm53148	Propionib
485	20	95.2	151	6	ABU38015	Abu38015	Protein e
486	20	95.2	151	6	ABU37100	Abu37100	Protein e
487	20	95.2	151	8	ADY09612	Ady09612	Plant ful
488	20	95.2	152	5	ABP26065	Abp26065	Streptoco
489	20	95.2	152	6	ABM15891	Abm15891	Mycobacte
490	20	95.2	153	3	AAB57110	Aab57110	Human pro
491	20	95.2	154	3	AAG40433	Aag40433	Arabidops
492	20	95.2	154	5	ABP64628	Abp64628	Human ORF
493	20	95.2	154	8	ADX92363	Adx92363	Plant ful
494	20	95.2	155	4	AAU56468	Aau56468	Propionib
495	20	95.2	155	6	ABM52987	Abm52987	Propionib
496	20	95.2	155	9	ABM95565	Abm95565	M. xanthu
497	20	95.2	156	4	AAM41931	Aam41931	Human pol
498	20	95.2	157	4	ABB70572	Abb70572	Drosophil
499	20	95.2	157	6	ABP79900	Abp79900	N. gonorr
500	20	95.2	158	3	AAG21092	Aag21092	Arabidops
501	20	95.2	158	3	AAG05681	Aag05681	Arabidops
502	20	95.2	158	3	AAG25206	Aag25206	Arabidops
503	20	95.2	158	4	AAU49514	Aau49514	Propionib
504	20	95.2	158	6	ABM46033	Abm46033	Propionib
505	20	95.2	158	6	ABM65700	Abm65700	Propionib
506	20	95.2	158	8	ADI13492	Adi13492	Linking r
507	20	95.2	158	8	ADK47131	Adk47131	Streptoco
508	20	95.2	158	8	ADR95100	Adr95100	Novel S.
509	20	95.2	158	8	ADV89708	Adv89708	Streptoco
510	20	95.2	158	8	ADV80961	Adv80961	Streptoco
511	20	95.2	158	8	ADX88039	Adx88039	Plant ful
512	20	95.2	158	9	AEA58970	Aea58970	Streptoco
513	20	95.2	159	2	AAY60121	Aay60121	Human end
514	20	95.2	159	4	AAU63809	Aau63809	Propionib
515	20	95.2	159	6	ABM60328	Abm60328	Propionib
516	20	95.2	160	4	AAG75051	Aag75051	Human col
517	20	95.2	160	5	AAE15947	Aae15947	Fly CP334
518	20	95.2	160	5	AAE15946	Aae15946	Fly CP334
519	20	95.2	160	8	ADQ16006	Adq16006	Human Mas
520	20	95.2	161	2	AAY48359	Aay48359	Human pro
521	20	95.2	162	3	AAG40432	Aag40432	Arabidops
522	20	95.2	162	8	ADS15061	Ads15061	Pseudomon
523	20	95.2	164	7	ADM04730	Adm04730	Human pro
524	20	95.2	165	2	AAW62673	Aaw62673	Streptoco
525	20	95.2	165	3	AAG31962	Aag31962	Arabidops
526	20	95.2	166	6	ABU02700	Abu02700	S. pneumo
527	20	95.2	166	8	ADK47792	Adk47792	Streptoco
528	20	95.2	167	4	ABB66864	Abb66864	Drosophil
529	20	95.2	167	4	ABB71479	Abb71479	Drosophil
530	20	95.2	167	4	ABG24639	Abg24639	Novel hum
531	20	95.2	167	4	ABG30245	Abg30245	Novel hum
532	20	95.2	167	7	ADC86897	Adc86897	Human GPC
533	20	95.2	168	8	ADR95678	Adr95678	Novel S.
534	20	95.2	168	8	ADY12762	Ady12762	Plant ful
535	20	95.2	168	9	AEA59548	Aea59548	Streptoco

536	20	95.2	169	7	ABO72175	Abo72175	Pseudomon
537	20	95.2	169	8	ADN26043	Adn26043	Bacterial
538	20	95.2	170	3	AAB25298	Aab25298	Eucalyptu
539	20	95.2	170	4	AAU62022	Aau62022	Propionib
540	20	95.2	170	6	ABM58541	Abm58541	Propionib
541	20	95.2	170	7	ABO74371	Abo74371	Pseudomon
542	20	95.2	170	7	ABO69833	Abo69833	Pseudomon
543	20	95.2	170	8	ADR94851	Adr94851	Novel S.
544	20	95.2	170	9	AEA58721	Aea58721	Streptoco
545	20	95.2	172	7	ABO73955	Abo73955	Pseudomon
546	20	95.2	172	8	ADY12741	Ady12741	Plant ful
547	20	95.2	172	9	ABM90696	Abm90696	M. xanthu
548	20	95.2	172	9	ABM93767	Abm93767	M. xanthu
549	20	95.2	173	4	ABG11981	Abg11981	Novel hum
550	20	95.2	173	6	ADA33948	Ada33948	Acinetoba
551	20	95.2	173	8	ADH22520	Adh22520	Human tra
552	20	95.2	174	5	AAU78275	Aau78275	Human rap
553	20	95.2	175	4	AAM41718	Aam41718	Human pol
554	20	95.2	175	4	AAU27520	Aau27520	Human G-P
555	20	95.2	175	5	AAU99296	Aau99296	Pyrococcu
556	20	95.2	175	8	ADT60505	Adt60505	Plant pol
557	20	95.2	176	8	ADT57037	Adt57037	Plant pol
558	20	95.2	176	8	ADX70906	Adx70906	Plant ful
559	20	95.2	177	6	ABU23025	Abu23025	Protein e
560	20	95.2	178	7	ABO79964	Abo79964	Pseudomon
561	20	95.2	179	3	AAG32991	Aag32991	Arabidops
562	20	95.2	179	8	ADV83107	Adv83107	Streptoco
563	20	95.2	180	3	AAB24670	Aab24670	Plant SDF
564	20	95.2	180	3	AAB24658	Aab24658	Plant SDF
565	20	95.2	180	3	AAG59347	Aag59347	Arabidops
566	20	95.2	180	7	ADC87059	Adc87059	Human GPC
567	20	95.2	180	7	ADF07384	Adf07384	Bacterial
568	20	95.2	180	8	ADU17867	Adu17867	Reverse t
569	20	95.2	180	8	ADU17864	Adu17864	Reverse t
570	20	95.2	180	9	ADY27674	Ady27674	Human ach
571	20	95.2	182	4	ABG16757	Abg16757	Novel hum
572	20	95.2	182	7	ABO71799	Abo71799	Pseudomon
573	20	95.2	183	8	ADL05999	Adl05999	M. catarr
574	20	95.2	183	8	ADX67620	Adx67620	Plant ful
575	20	95.2	183	8	ADX78430	Adx78430	Plant ful
576	20	95.2	184	4	AAB64436	Aab64436	Human sec
577	20	95.2	185	4	ABG19120	Abg19120	Novel hum
578	20	95.2	185	7	ADM25943	Adm25943	Hyperther
579	20	95.2	185	7	ABO62926	Abo62926	Klebsiell
580	20	95.2	185	8	ADL06062	Adl06062	M. catarr
581	20	95.2	185	8	ADX88139	Adx88139	Plant ful
582	20	95.2	187	2	AAY38513	Aay38513	Neisseria
583	20	95.2	187	3	AAB03148	Aab03148	Human neu
584	20	95.2	187	9	ABM93390	Abm93390	M. xanthu
585	20	95.2	187	9	ABM91093	Abm91093	M. xanthu
586	20	95.2	187	9	AEB48883	Aeb48883	N. mening
587	20	95.2	188	7	ADF07023	Adf07023	Bacterial
588	20	95.2	188	7	ADM25787	Adm25787	Hyperther
589	20	95.2	188	8	ABO59539	Abo59539	Human gen
590	20	95.2	188	8	ADY07159	Ady07159	Plant ful
591	20	95.2	189	3	AAB43607	Aab43607	Human can
592	20	95.2	189	3	AAG60392	Aag60392	Arabidops
593	20	95.2	189	3	AAG57484	Aag57484	Arabidops
594	20	95.2	189	7	ADD18799	Add18799	Human dis
595	20	95.2	189	8	ABO59519	Abo59519	Human gen
596	20	95.2	189	8	ADS88270	Ads88270	Human pro
597	20	95.2	189	9	ADY18038	Ady18038	PRO polyp
598	20	95.2	190	3	AAG32452	Aag32452	Arabidops
599	20	95.2	190	3	AAG26346	Aag26346	Arabidops
600	20	95.2	190	4	ABB70193	Abb70193	Drosophil
601	20	95.2	190	7	ABO80877	Abo80877	Pseudomon
602	20	95.2	191	4	ABB66296	Abb66296	Drosophil
603	20	95.2	191	8	ADY09572	Ady09572	Plant ful
604	20	95.2	192	2	AAW64369	Aaw64369	Mycobacte
605	20	95.2	192	2	AAW81734	Aaw81734	M. tuberc
606	20	95.2	192	2	AAY39023	Aay39023	M. tuberc
607	20	95.2	192	2	AAY39166	Aay39166	M. tuberc
608	20	95.2	192	3	AAG26192	Aag26192	Arabidops
609	20	95.2	192	3	AAG28658	Aag28658	Arabidops
610	20	95.2	192	3	AAG53689	Aag53689	Arabidops
611	20	95.2	192	4	AAU45522	Aau45522	Propionib
612	20	95.2	192	4	ABG23080	Abg23080	Novel hum
613	20	95.2	192	6	ABM42041	Abm42041	Propionib
614	20	95.2	192	8	ADR20249	Adr20249	Herv-K/HM
615	20	95.2	192	8	ADR95795	Adr95795	Novel S.
616	20	95.2	192	8	ADT56323	Adt56323	Plant pol
617	20	95.2	192	9	ADW72805	Adw72805	Human NKp
618	20	95.2	192	9	AEA40203	Aea40203	Human nat
619	20	95.2	192	9	AEA59665	Aea59665	Streptoco
620	20	95.2	193	4	ABB66326	Abb66326	Drosophil
621	20	95.2	193	7	ABO75621	Abo75621	Pseudomon
622	20	95.2	193	9	ABM93373	Abm93373	M. xanthu
623	20	95.2	194	3	AAG08208	Aag08208	Arabidops
624	20	95.2	194	5	ABB08622	Abb08622	Human DNA
625	20	95.2	195	8	ADT57977	Adt57977	Plant pol
626	20	95.2	196	4	AAG90590	Aag90590	C glutami
627	20	95.2	196	4	ABG11745	Abg11745	Novel hum
628	20	95.2	196	6	ABR41472	Abr41472	Human DIT
629	20	95.2	197	4	ABG16546	Abg16546	Novel hum
630	20	95.2	198	3	AAG01347	Aag01347	Human sec
631	20	95.2	199	7	ADM26103	Adm26103	Hyperther
632	20	95.2	199	7	ABO74256	Abo74256	Pseudomon
633	20	95.2	199	8	ADU17770	Adu17770	Reverse t
634	20	95.2	199	8	ADU17793	Adu17793	Reverse t
635	20	95.2	199	8	ADU17783	Adu17783	Reverse t
636	20	95.2	199	8	ADU17794	Adu17794	Reverse t
637	20	95.2	199	8	ADU17767	Adu17767	Reverse t
638	20	95.2	199	8	ADU17768	Adu17768	Reverse t
639	20	95.2	199	8	ADU17766	Adu17766	Reverse t
640	20	95.2	199	8	ADU17772	Adu17772	Reverse t
641	20	95.2	199	8	ADU17795	Adu17795	Reverse t
642	20	95.2	200	2	AAW28317	Aaw28317	Amino aci
643	20	95.2	201	3	AAG21619	Aag21619	Arabidops
644	20	95.2	201	4	AAB63738	Aab63738	Human pro
645	20	95.2	201	7	ABO79454	Abo79454	Pseudomon
646	20	95.2	202	5	ABB90006	Abb90006	Human pol
647	20	95.2	202	7	ABO73570	Abo73570	Pseudomon
648	20	95.2	203	4	AAB93444	Aab93444	Human pro
649	20	95.2	203	5	ABB89562	Abb89562	Human pol
650	20	95.2	203	6	ABU56465	Abu56465	Lung canc
651	20	95.2	203	8	ADH18880	Adh18880	Human cel
652	20	95.2	203	8	ADN05076	Adn05076	Antipsori
653	20	95.2	205	4	AAU67257	Aau67257	Propionib
654	20	95.2	205	5	ABB94341	Abb94341	Chlamydia
655	20	95.2	205	5	ABU05791	Abu05791	M. tuberc
656	20	95.2	205	5	ABU05386	Abu05386	M. tuberc
657	20	95.2	205	6	ABM63776	Abm63776	Propionib
658	20	95.2	205	7	ABO76651	Abo76651	Pseudomon
659	20	95.2	206	2	AAR81881	Aar81881	Mouse typ
660	20	95.2	206	2	AAW48977	Aaw48977	Mouse OX4
661	20	95.2	207	8	ADJ48873	Adj48873	Oil-assoc
662	20	95.2	208	5	ABP38028	Abp38028	Staphyloc
663	20	95.2	208	6	ABR52585	Abr52585	Protein s
664	20	95.2	208	7	ADK63902	Adk63902	Disease t
665	20	95.2	208	7	ABO82964	Abo82964	Pseudomon
666	20	95.2	209	4	ABB68656	Abb68656	Drosophil
667	20	95.2	209	4	ABG15663	Abg15663	Novel hum
668	20	95.2	209	9	ADW72804	Adw72804	Human NKp
669	20	95.2	209	9	AEA40202	Aea40202	Human nat
670	20	95.2	210	4	AAB63739	Aab63739	Human pro
671	20	95.2	210	4	AAB63742	Aab63742	Human pro
672	20	95.2	210	5	ABU05448	Abu05448	M. tuberc
673	20	95.2	210	6	ABU41414	Abu41414	Protein e
674	20	95.2	210	8	ADT57410	Adt57410	Plant pol
675	20	95.2	210	8	ADX97087	Adx97087	Plant ful
676	20	95.2	211	4	AAM18601	Aam18601	Peptide #
677	20	95.2	211	4	ABB37657	Abb37657	Peptide #
678	20	95.2	211	4	AAM31058	Aam31058	Peptide #
679	20	95.2	211	4	ABB32378	Abb32378	Peptide #
680	20	95.2	211	4	AAM39932	Aam39932	Human pol
681	20	95.2	211	4	AAM58289	Aam58289	Human bra

682	20	95.2	211	4	ABG52465	Abg52465	Human liv
683	20	95.2	211	4	AAM06172	Aam06172	Peptide #
684	20	95.2	211	5	ABG40501	Abg40501	Human pep
685	20	95.2	211	8	ADX91661	Adx91661	Plant ful
686	20	95.2	212	2	AAY35210	Aay35210	Chlamydia
687	20	95.2	212	3	AAG21618	Aag21618	Arabidops
688	20	95.2	212	4	AAU19495	Aau19495	Human dia
689	20	95.2	212	4	AAU67171	Aau67171	Propionib
690	20	95.2	212	4	ABG05041	Abg05041	Novel hum
691	20	95.2	212	6	ABM63690	Abm63690	Propionib
692	20	95.2	212	7	ABO77775	Abo77775	Pseudomon
693	20	95.2	213	4	AAU19186	Aau19186	Human G p
694	20	95.2	213	4	AAU50790	Aau50790	Propionib
695	20	95.2	213	6	ABU02007	Abu02007	S. pneumo
696	20	95.2	213	6	ABM47309	Abm47309	Propionib
697	20	95.2	213	7	ADF59061	Adf59061	Human pol
698	20	95.2	213	7	ABO82012	Abo82012	Pseudomon
699	20	95.2	214	4	AAU47428	Aau47428	Propionib
700	20	95.2	214	4	ABG00921	Abg00921	Novel hum
701	20	95.2	214	6	AAE32354	Aae32354	Murine FD
702	20	95.2	214	6	ABM43947	Abm43947	Propionib
703	20	95.2	214	7	ADD25201	Add25201	Fertility
704	20	95.2	214	7	ADD69285	Add69285	Murine FD
705	20	95.2	214	8	ADN61216	Adn61216	Radish nu
706	20	95.2	215	4	AAO01838	Aao01838	Human pol
707	20	95.2	215	5	ABP30487	Abp30487	Streptoco
708	20	95.2	215	8	ADU69658	Adu69658	S agalact
709	20	95.2	215	8	ADV88193	Adv88193	Streptoco
710	20	95.2	215	8	ADV79446	Adv79446	Streptoco
711	20	95.2	215	8	ADV81621	Adv81621	Streptoco
712	20	95.2	216	4	ABU52791	Abu52791	Human sig
713	20	95.2	216	5	ABG79610	Abg79610	Para rubb
714	20	95.2	216	8	ADY09061	Ady09061	Plant ful
715	20	95.2	217	5	AAE13525	Aae13525	Mouse FDR
716	20	95.2	218	3	AAG41655	Aag41655	Arabidops
717	20	95.2	219	4	AAU23415	Aau23415	Novel hum
718	20	95.2	220	3	ABB08911	Abb08911	Streptomy
719	20	95.2	220	4	AAG73673	Aag73673	Human col
720	20	95.2	220	6	ABM65628	Abm65628	Propionib
721	20	95.2	220	8	ADS24417	Ads24417	Bacterial
722	20	95.2	221	4	ABB60165	Abb60165	Drosophil
723	20	95.2	222	8	ADX89423	Adx89423	Plant ful
724	20	95.2	222	8	ADX67284	Adx67284	Plant ful
725	20	95.2	223	8	ADX68786	Adx68786	Plant ful
726	20	95.2	224	5	AAO21843	Aao21843	Pathogeni
727	20	95.2	224	5	ABB09133	Abb09133	Thermotol
728	20	95.2	224	5	ABB98517	Abb98517	Pyrococcu
729	20	95.2	224	5	ABP53222	Abp53222	Pyrococcu
730	20	95.2	224	5	AAU78000	Aau78000	Pyrococcu
731	20	95.2	224	5	ABB97745	Abb97745	P. furios
732	20	95.2	224	6	ABM69759	Abm69759	Photorhab
733	20	95.2	224	7	ADC55436	Adc55436	Protein s
734	20	95.2	224	8	ADR99391	Adr99391	Chlamydia
735	20	95.2	224	9	AEB40994	Aeb40994	L. pneumo
736	20	95.2	224	9	AEB37683	Aeb37683	L. pneumo
737	20	95.2	225	3	AAG25205	Aag25205	Arabidops
738	20	95.2	225	9	ABM90636	Abm90636	M. xanthu
739	20	95.2	227	7	ADF05977	Adf05977	Bacterial
740	20	95.2	227	8	ADU02756	Adu02756	Novel hum
741	20	95.2	228	3	AAG40130	Aag40130	Arabidops
742	20	95.2	228	3	AAG07467	Aag07467	Arabidops
743	20	95.2	228	4	ABG18546	Abg18546	Novel hum
744	20	95.2	229	3	AAG26191	Aag26191	Arabidops
745	20	95.2	229	4	ABG14362	Abg14362	Novel hum
746	20	95.2	230	3	AAG57483	Aag57483	Arabidops
747	20	95.2	230	3	AAG60391	Aag60391	Arabidops
748	20	95.2	230	8	ADO17017	Ado17017	Predicted
749	20	95.2	230	8	ADS29937	Ads29937	Bacterial
750	20	95.2	231	2	AAW60988	Aaw60988	Streptoco
751	20	95.2	231	3	AAB13459	Aab13459	Protein e
752	20	95.2	231	3	AAG49381	Aag49381	Arabidops
753	20	95.2	231	3	AAG41654	Aag41654	Arabidops
754	20	95.2	231	4	AAG83135	Aag83135	Protein e
755	20	95.2	231	5	ABB94106	Abb94106	Chlamydia
756	20	95.2	231	5	AAE18736	Aae18736	Human lun
757	20	95.2	231	6	ABM71291	Abm71291	Staphyloc
758	20	95.2	231	8	ADY05007	Ady05007	Plant ful
759	20	95.2	232	8	ADK70536	Adk70536	Respirato
760	20	95.2	233	3	AAY74686	Aay74686	Neisseria
761	20	95.2	233	4	ABG24475	Abg24475	Novel hum
762	20	95.2	233	4	ABG29906	Abg29906	Novel hum
763	20	95.2	233	5	ABP27713	Abp27713	Streptoco
764	20	95.2	233	5	AAE18735	Aae18735	Human lun
765	20	95.2	233	6	ABP75496	Abp75496	Human sec
766	20	95.2	233	6	ABU06043	Abu06043	N. lactam
767	20	95.2	233	7	ADB64902	Adb64902	Human pro
768	20	95.2	233	7	ADF59565	Adf59565	Human pol
769	20	95.2	233	7	ADF60565	Adf60565	Human con
770	20	95.2	233	7	ADI21122	Adi21122	Novel hum
771	20	95.2	234	2	AAY59835	Aay59835	Human nor
772	20	95.2	234	4	ABG21806	Abg21806	Novel hum
773	20	95.2	234	5	AAU85594	Aau85594	Lung tumo
774	20	95.2	234	5	ABB97245	Abb97245	Novel hum
775	20	95.2	234	6	ABU02699	Abu02699	S. pneumo
776	20	95.2	234	6	ABU69566	Abu69566	Human lun
777	20	95.2	234	6	ABP81497	Abp81497	Streptoco
778	20	95.2	234	6	ABU66469	Abu66469	Lung canc
779	20	95.2	234	7	ADH47391	Adh47391	Human lun
780	20	95.2	234	8	ADK47790	Adk47790	Streptoco
781	20	95.2	234	8	ADJ21310	Adj21310	Human lun
782	20	95.2	234	9	AEB03554	Aeb03554	Mycobacte
783	20	95.2	234	9	AEA79457	Aea79457	Novel M.
784	20	95.2	235	2	AAR70101	Aar70101	LIYV RNA
785	20	95.2	235	3	AAG19255	Aag19255	Arabidops
786	20	95.2	235	3	AAG43087	Aag43087	Arabidops
787	20	95.2	235	4	AAU46627	Aau46627	Propionib
788	20	95.2	235	6	ABM43146	Abm43146	Propionib
789	20	95.2	236	2	AAY38516	Aay38516	Neisseria
790	20	95.2	236	4	ABG22704	Abg22704	Novel hum
791	20	95.2	236	5	AAU85592	Aau85592	Lung tumo
792	20	95.2	236	5	AAU85591	Aau85591	Lung tumo
793	20	95.2	236	5	AAE29287	Aae29287	Human mas
794	20	95.2	236	6	ADA83740	Ada83740	Human ASC
795	20	95.2	236	6	ABU69563	Abu69563	Human lun
796	20	95.2	236	6	ABU69564	Abu69564	Human lun
797	20	95.2	236	6	ABU66466	Abu66466	Lung canc
798	20	95.2	236	6	ABU66467	Abu66467	Lung canc
799	20	95.2	236	6	ADB11136	Adb11136	Alloiococ
800	20	95.2	236	7	ADH47389	Adh47389	Human lun
801	20	95.2	236	7	ADH47388	Adh47388	Human lun
802	20	95.2	236	7	ABO82278	Abo82278	Pseudomon
803	20	95.2	236	8	ADQ16003	Adq16003	Human Mas
804	20	95.2	236	8	ADJ21308	Adj21308	Human lun
805	20	95.2	236	8	ADJ21307	Adj21307	Human lun
806	20	95.2	236	8	ABM81462	Abm81462	Tumour-as
807	20	95.2	236	8	ADS84968	Ads84968	Human lun
808	20	95.2	236	9	ADY49491	Ady49491	Human Mas
809	20	95.2	236	9	ADY27671	Ady27671	Human ach
810	20	95.2	236	9	AEB48888	Aeb48888	N. gonorr
811	20	95.2	237	5	AAU85593	Aau85593	Lung tumo
812	20	95.2	237	6	ABR40861	Abr40861	Zea mays
813	20	95.2	237	6	ABU69565	Abu69565	Human lun
814	20	95.2	237	6	ABU66468	Abu66468	Lung canc
815	20	95.2	237	7	ADH47390	Adh47390	Human lun
816	20	95.2	237	8	ADJ21309	Adj21309	Human lun
817	20	95.2	238	3	AAB14348	Aab14348	Human ASH
818	20	95.2	238	4	AAU33967	Aau33967	Staphyloc
819	20	95.2	238	4	ABG15154	Abg15154	Novel hum
820	20	95.2	238	5	AAE18734	Aae18734	Human lun
821	20	95.2	238	5	AAU85551	Aau85551	Clone #50
822	20	95.2	238	5	ABU05365	Abu05365	M. tuberc
823	20	95.2	238	5	AAE29288	Aae29288	Human mas
824	20	95.2	238	6	ABU56502	Abu56502	Lung canc
825	20	95.2	238	6	ABU69523	Abu69523	Human lun
826	20	95.2	238	6	ABU66426	Abu66426	Lung canc
827	20	95.2	238	7	ADB74503	Adb74503	Mycobacte

828	20	95.2	238	7	ADH47332	Adh47332	Human lun	901	20	95.2	255	3	AAY74711	Aay74711	Neisseria
829	20	95.2	238	8	ADJ21251	Adj21251	Human lun	902	20	95.2	255	3	AAG21617	Aag21617	Arabidops
830	20	95.2	238	9	ADV70192	Adv70192	Tumor-ass	903	20	95.2	255	3	AAG21660	Aag21660	Arabidops
831	20	95.2	238	9	ADY49493	Ady49493	Human Mas	904	20	95.2	255	3	AAG52164	Aag52164	Arabidops
832	20	95.2	239	2	AAW80690	Aaw80690	S. pneumo	905	20	95.2	255	4	AAB92574	Aab92574	Human pro
833	20	95.2	239	3	AAG07840	Aag07840	Arabidops	906	20	95.2	255	4	AAM93970	Aam93970	Human sto
834	20	95.2	239	8	ADG93991	Adg93991	Escherich	907	20	95.2	255	6	ABP79581	Abp79581	N. gonorr
835	20	95.2	239	8	ADJ54176	Adj54176	MarA tran	908	20	95.2	255	8	ADT49856	Adt49856	Murine FB
836	20	95.2	239	8	ADO15500	Ado15500	Escherich	909	20	95.2	255	8	ADY22900	Ady22900	Plant ful
837	20	95.2	239	9	AEA32579	Aea32579	Escherich	910	20	95.2	256	4	ABB61922	Abb61922	Drosophil
838	20	95.2	240	4	AAU37860	Aau37860	Streptoco	911	20	95.2	256	8	ADN73617	Adn73617	Thale cre
839	20	95.2	240	4	AAM40842	Aam40842	Human pol	912	20	95.2	257	2	AAW56493	Aaw56493	Tobacco l
840	20	95.2	240	5	ABP27775	Abp27775	Streptoco	913	20	95.2	257	4	ABG29649	Abg29649	Novel hum
841	20	95.2	240	6	ABU02367	Abu02367	S. pneumo	914	20	95.2	257	6	ABU48422	Abu48422	Protein e
842	20	95.2	240	6	ABU46228	Abu46228	Protein e	915	20	95.2	257	7	ADH88448	Adh88448	Enterococ
843	20	95.2	240	6	ABU46476	Abu46476	Protein e	916	20	95.2	257	9	ADV09395	Adv09395	TYLCSV ca
844	20	95.2	240	6	ABU44230	Abu44230	Protein e	917	20	95.2	258	4	AAM41766	Aam41766	Human pol
845	20	95.2	240	8	ADK46455	Adk46455	Streptoco	918	20	95.2	258	6	ABM67981	Abm67981	Photorhab
846	20	95.2	240	8	ADV88518	Adv88518	Streptoco	919	20	95.2	258	8	ADY13339	Ady13339	Plant ful
847	20	95.2	240	8	ADV79771	Adv79771	Streptoco	920	20	95.2	259	5	ABG77446	Abg77446	Selected
848	20	95.2	240	8	ADV81928	Adv81928	Streptoco	921	20	95.2	259	5	ABB54718	Abb54718	Lactococc
849	20	95.2	241	4	AAU38016	Aau38016	Streptoco	922	20	95.2	259	5	ABB54375	Abb54375	Lactococc
850	20	95.2	241	6	ABR52784	Abr52784	Protein s	923	20	95.2	259	5	ABB53792	Abb53792	Lactococc
851	20	95.2	241	7	ADK62144	Adk62144	Disease t	924	20	95.2	259	5	ABJ11348	Abj11348	Yeast sel
852	20	95.2	241	7	ABO64120	Abo64120	Klebsiell	925	20	95.2	259	7	ABM87612	Abm87612	Rice abio
853	20	95.2	241	8	ADQ65226	Adq65226	Novel hum	926	20	95.2	260	4	AAU35725	Aau35725	Helicobac
854	20	95.2	242	3	AAG23135	Aag23135	Arabidops	927	20	95.2	260	4	AAG79123	Aag79123	Amino aci
855	20	95.2	242	4	AAU36571	Aau36571	Staphyloc	928	20	95.2	260	6	ABR64003	Abr64003	Predicted
856	20	95.2	242	5	ABB54066	Abb54066	Lactococc	929	20	95.2	260	6	ABU30762	Abu30762	Protein e
857	20	95.2	242	6	ABU15970	Abu15970	Protein e	930	20	95.2	260	8	ADQ48612	Adq48612	Helicobac
858	20	95.2	242	6	ABU23340	Abu23340	Protein e	931	20	95.2	260	8	ADS30639	Ads30639	Bacterial
859	20	95.2	242	8	ADK46819	Adk46819	Streptoco	932	20	95.2	261	4	ABB68328	Abb68328	Drosophil
860	20	95.2	242	8	ADR95703	Adr95703	Novel S.	933	20	95.2	261	4	AAG81584	Aag81584	S. epider
861	20	95.2	242	9	AEA59573	Aea59573	Streptoco	934	20	95.2	261	4	AAG82674	Aag82674	S. epider
862	20	95.2	244	4	ABG26006	Abg26006	Novel hum	935	20	95.2	261	8	ADL04383	Adl04383	M. catarr
863	20	95.2	244	6	ABR40863	Abr40863	Zea mays	936	20	95.2	262	3	AAG07839	Aag07839	Arabidops
864	20	95.2	244	8	ADT57497	Adt57497	Plant pol	937	20	95.2	262	7	ADI21247	Adi21247	Novel hum
865	20	95.2	245	8	ABO58572	Abo58572	Human gen	938	20	95.2	263	4	AAU17415	Aau17415	Novel sig
866	20	95.2	245	8	ADN24505	Adn24505	Bacterial	939	20	95.2	263	6	ABU52589	Abu52589	Human NOV
867	20	95.2	246	3	AAG25204	Aag25204	Arabidops	940	20	95.2	263	7	ADB94123	Adb94123	Human nov
868	20	95.2	246	6	ABR40673	Abr40673	Zea mays	941	20	95.2	263	8	ADS21535	Ads21535	Bacterial
869	20	95.2	246	9	ABM93456	Abm93456	M. xanthu	942	20	95.2	264	5	ABB47511	Abb47511	Listeria
870	20	95.2	247	3	AAG22404	Aag22404	Arabidops	943	20	95.2	264	8	ADU02417	Adu02417	Novel hum
871	20	95.2	247	3	AAG12671	Aag12671	Arabidops	944	20	95.2	266	2	AAY43940	Aay43940	Human pro
872	20	95.2	247	4	ABU53320	Abu53320	Human cel	945	20	95.2	266	6	ABU45250	Abu45250	Protein e
873	20	95.2	247	4	ABU52991	Abu52991	Human sig	946	20	95.2	267	6	ABU25437	Abu25437	Protein e
874	20	95.2	247	6	ABJ25658	Abj25658	Aspergill	947	20	95.2	267	8	ADJ50291	Adj50291	Oil-assoc
875	20	95.2	247	8	ADX89842	Adx89842	Plant ful	948	20	95.2	267	8	ADN27291	Adn27291	Bacterial
876	20	95.2	247	9	ADW18364	Adw18364	Pinus rad	949	20	95.2	267	8	ADX68014	Adx68014	Plant ful
877	20	95.2	248	2	AAY35303	Aay35303	Chlamydia	950	20	95.2	267	9	ABM95402	Abm95402	M. xanthu
878	20	95.2	248	3	AAY75766	Aay75766	Neisseria	951	20	95.2	269	4	ABB63346	Abb63346	Drosophil
879	20	95.2	248	3	AAY75767	Aay75767	Neisseria	952	20	95.2	269	4	AAM40145	Aam40145	Human pol
880	20	95.2	248	3	AAY75765	Aay75765	Neisseria	953	20	95.2	269	5	AAU80184	Aau80184	Embryonic
881	20	95.2	248	5	ABP27776	Abp27776	Streptoco	954	20	95.2	269	8	ADR09206	Adr09206	Human pro
882	20	95.2	248	6	ABM71601	Abm71601	Staphyloc	955	20	95.2	270	4	AAG82813	Aag82813	S. epider
883	20	95.2	248	7	ADM26586	Adm26586	Hypether	956	20	95.2	270	6	ABR41385	Abr41385	Human DIT
884	20	95.2	248	8	ADS43965	Ads43965	Bacterial	957	20	95.2	272	2	AAY29201	Aay29201	Amino aci
885	20	95.2	250	8	ADX88498	Adx88498	Plant ful	958	20	95.2	272	3	AAB58955	Aab58955	Breast an
886	20	95.2	250	8	ADX89753	Adx89753	Plant ful	959	20	95.2	272	3	AAG31961	Aag31961	Arabidops
887	20	95.2	251	3	AAG52165	Aag52165	Arabidops	960	20	95.2	272	5	ABG78032	Abg78032	ITALY, LO
888	20	95.2	251	3	AAG21661	Aag21661	Arabidops	961	20	95.2	272	8	ADR45558	Adr45558	Homologue
889	20	95.2	251	8	ADN99247	Adn99247	Novel hum	962	20	95.2	273	3	AAG17674	Aag17674	Arabidops
890	20	95.2	251	8	ADX75768	Adx75768	Plant ful	963	20	95.2	273	3	AAG47587	Aag47587	Arabidops
891	20	95.2	252	3	AAG57482	Aag57482	Arabidops	964	20	95.2	273	4	AAU58254	Aau58254	Propionib
892	20	95.2	252	3	AAG60390	Aag60390	Arabidops	965	20	95.2	273	5	ABB93858	Abb93858	Herbicida
893	20	95.2	252	4	AAM79289	Aam79289	Human pro	966	20	95.2	273	6	ABM54773	Abm54773	Propionib
894	20	95.2	252	7	ABR39143	Abr39143	TRV-RNA1	967	20	95.2	273	6	ABU06044	Abu06044	N. mening
895	20	95.2	252	7	ADM25408	Adm25408	Hyperther	968	20	95.2	273	6	ABU37768	Abu37768	Protein e
896	20	95.2	253	3	AAY71553	Aay71553	Soybean s	969	20	95.2	274	2	AAY17979	Aay17979	Hexulose
897	20	95.2	253	7	ABO66960	Abo66960	Klebsiell	970	20	95.2	275	2	AAY60029	Aay60029	Human end
898	20	95.2	253	8	ADG22675	Adg22675	Cyanophag	971	20	95.2	275	8	ADN24570	Adn24570	Bacterial
899	20	95.2	254	7	ADF58819	Adf58819	Human pol	972	20	95.2	275	8	ADN21810	Adn21810	Bacterial
900	20	95.2	254	8	ADN47979	Adn47979	Thermococ	973	20	95.2	276	3	AAG14069	Aag14069	Arabidops

974	20	95.2	276	5	ABB91124	Abb91124 Herbicida
975	20	95.2	276	6	ADA34395	Ada34395 Acinetoba
976	20	95.2	276	8	ADX71305	Adx71305 Plant ful
977	20	95.2	277	3	AAG30215	Aag30215 Arabidops
978	20	95.2	277	4	ABG04213	Abg04213 Novel hum
979	20	95.2	278	6	ABM70140	Abm70140 Photorhab
980	20	95.2	278	7	ABO66293	Abo66293 Klebsiell
981	20	95.2	279	2	AAR91312	Aar91312 N. gonorr
982	20	95.2	279	6	ABP77390	Abp77390 N. gonorr
983	20	95.2	279	6	ABU50637	Abu50637 Protein e
984	20	95.2	280	3	AAG08396	Aag08396 Arabidops
985	20	95.2	280	7	ABO65870	Abo65870 Klebsiell
986	20	95.2	281	7	ABO79112	Abo79112 Pseudomon
987	20	95.2	281	8	ADY09043	Ady09043 Plant ful
988	20	95.2	282	2	AAY48330	Aay48330 Human pro
989	20	95.2	282	3	AAG30214	Aag30214 Arabidops
990	20	95.2	282	5	ABB92302	Abb92302 Herbicida
991	20	95.2	282	6	ABP79606	Abp79606 N. gonorr
992	20	95.2	282	7	ADE12782	Ade12782 L. rhamno
993	20	95.2	282	8	ADS28436	Ads28436 Bacterial
994	20	95.2	283	3	AAG14068	Aag14068 Arabidops
995	20	95.2	283	7	ADC86377	Adc86377 Human GPC
996	20	95.2	284	6	ABU44340	Abu44340 Protein e
997	20	95.2	284	8	ADN18215	Adn18215 Bacterial
998	20	95.2	285	5	ABB90597	Abb90597 Chlamydia
999	20	95.2	285	7	ABO67704	Abo67704 Klebsiell
1000	20	95.2	285	8	ADJ48372	Adj48372 Maize oil

ALIGNMENTS

RESULT 1
ADZ71770
ID ADZ71770 standard; peptide; 5 AA.
XX
AC ADZ71770;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #355.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 22; Page 103; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

  Query Match   95.2%; Score 20; DB 9; Length 5;
  Best Local Similarity   100.0%; Pred. No. 2e+06;
  Matches   4; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   1 RRLN 4
              ||||
Db   1 RRLN 4

RESULT 2
ADZ72025
ID ADZ72025 standard; peptide; 5 AA.
XX
AC ADZ72025;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #610.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX

PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Example 27; Page 83; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 3
ADZ71763
ID ADZ71763 standard; peptide; 5 AA.
XX
AC ADZ71763;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #348.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 22; Page 103; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay

CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 4
ADZ72027
ID ADZ72027 standard; peptide; 5 AA.
XX
AC ADZ72027;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #612.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Example 27; Page 83; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 5
ADZ72107
ID ADZ72107 standard; peptide; 5 AA.
XX
AC ADZ72107;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #692.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.

XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Example 27; Page 85; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 6
ADZ71617
ID ADZ71617 standard; peptide; 5 AA.
XX
AC ADZ71617;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #202.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 15; Page 98; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate

CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 7
ADZ71663
ID ADZ71663 standard; peptide; 5 AA.
XX
AC ADZ71663;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #248.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 20; Page 100; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 8
ADZ71764
ID ADZ71764 standard; peptide; 5 AA.
XX
AC ADZ71764;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #349.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX

PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 22; Page 103; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 9
ADZ71765
ID ADZ71765 standard; peptide; 5 AA.
XX
AC ADZ71765;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #350.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 22; Page 103; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.

CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 10
ADZ71746
ID ADZ71746 standard; peptide; 5 AA.
XX
AC ADZ71746;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #331.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 21; Page 102; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 11
ADZ71618
ID ADZ71618 standard; peptide; 5 AA.
XX
AC ADZ71618;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #203.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.

XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 15; Page 98; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from:
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 12
ADZ71662
ID ADZ71662 standard; peptide; 5 AA.
XX
AC ADZ71662;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #247.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 20; Page 100; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the

CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 13
ADZ71664
ID ADZ71664 standard; peptide; 5 AA.
XX
AC ADZ71664;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #249.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 20; Page 100; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 14
ADZ72026
ID ADZ72026 standard; peptide; 5 AA.
XX
AC ADZ72026;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #611.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX

PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Example 27; Page 83; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||

Db 1 RRLN 4

RESULT 15
ADZ72106
ID ADZ72106 standard; peptide; 5 AA.
XX
AC ADZ72106;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #691.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Example 27; Page 85; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the

CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 16
ADZ71747
ID ADZ71747 standard; peptide; 5 AA.
XX
AC ADZ71747;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #332.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 21; Page 102; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 17
ADZ71748
ID ADZ71748 standard; peptide; 5 AA.
XX
AC ADZ71748;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #333.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.

PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Claim 21; Page 102; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 5 AA;

Query Match 95.2%; Score 20; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 18
AAW56904
ID AAW56904 standard; peptide; 6 AA.
XX
AC AAW56904;
XX
DT 28-JUL-1998 (first entry)
XX
DE Enzyme inhibitor peptide SEQ ID NO:105.
XX
KW Enzyme inhibitor; t-PA; u-PA; chymotrypsin; serine protease; active;
KW latent; substrate subtraction phage display peptide library;
KW identification; kinase; phosphatase; serpin.
XX
OS Homo sapiens.
XX
PN WO9747314-A1.
XX
PD 18-DEC-1997.
XX
PF 10-JUN-1997; 97WO-US009760.
XX
PR 10-JUN-1996; 96US-0019495P.
XX
PA (SCRI) SCRIPPS RES INST.
XX
PI Madison EL, Ke S;
XX
DR WPI; 1998-062746/06.
XX
PT Substrate subtraction phage display peptide libraries - used to
PT distinguish between active and latent forms of enzyme, e.g. serine
PT protease.
XX
PS Claim 25; Page 75; 138pp; English.
XX
CC The present sequence represents an enzyme inhibitor peptide used in the
CC method of the invention to distinguish between t-PA and u-PA. The present
CC invention describes a substrate subtraction library for the
CC identification of peptide substrates selective between a first enzyme
CC (E1) and a second enzyme (E2), comprising a collection different
CC peptides, substantially lacking peptides that are effective substrates
CC for E1. Also described are: (1) a method (M1) for identifying peptide
CC substrates selective between a first enzyme (E1) and a second enzyme (E2)
CC ; (2) a compound comprising the amino acid sequence of a peptide
CC identified by M1; (3) a polypeptide for use as an enzyme inhibitor
CC comprising one of 237 amino acid sequences (see AAW56801 to AAW56947, and
CC AAW56949 to AAW57038); (4) a recombinant DNA vector comprising DNA (I)
CC encoding a protease inhibitor including the sequence identified by the M1
CC ; (5) a prokaryotic or eukaryotic cell containing the vector of (4); (6)
CC an antibody (Ab) immunoreactive with at least one of the peptides
CC identified by M1; and (7) a diagnostic assay for distinguishing between
CC active and latent forms of protease inhibitors, that uses (Ab). The
CC library and method are used for distinguishing between active and latent
CC forms of enzyme inhibitors, e.g. proteases, kinases and phosphatases.
CC (Ab) are used for affinity purification of recombinant peptides and in
CC the identification of naturally occurring protease inhibitors. Enzyme-
CC inhibiting peptides identified can be used to treat a serpin deficiency
CC or a disorder of serine proteases
XX
SQ Sequence 6 AA;

Query Match 95.2%; Score 20; DB 2; Length 6;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 2 RRLN 5

```
RESULT 19
AAM46084
ID   AAM46084 standard; peptide; 7 AA.
XX
AC   AAM46084;
XX
DT   25-OCT-2001  (first entry)
XX
DE   H11 binding site consensus conforming peptide (CCP) #2355.
XX
KW   Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW   immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW   cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW   astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW   ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   CA2290722-A1.
XX
PD   08-JUN-2001.
XX
PF   08-DEC-1999;   99CA-02290722.
XX
PR   08-DEC-1999;   99CA-02290722.
XX
PA   (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI   Kaplan HA,  Maiti PK,  Fast DG,  Herman W,  Dan MD,  Lewis KE;
PI   Entwistle JM,  Macdonald GC;
XX
DR   WPI; 2001-425937/46.
XX
PT   Composition useful for treating and diagnosing cancer, comprises stress
PT   protein-peptide complexes associated with tumor, and isolated antigen-
PT   binding fragments of an antibody that binds specifically to the complex.
XX
PS   Example 4; Page 109; 154pp; English.
XX
CC   The present invention describes a composition (I) comprising stress
CC   protein-peptide complexes (SPPC) associated with tumours that is
CC   specifically immunogenically cross-reactive with cell surface-associated
CC   SPPCs specific to target cancer (TC). Also described is an isolated
CC   antigen-binding fragment of an antibody that binds specifically to SPPCs
CC   or a population of different SPPCs consisting of immunogenic cancer cell
CC   surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC   used in vaccine production and as a tumour-specific immunogenic response
CC   inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC   subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC   oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC   ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC   vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC   or imaging cancer cells, and to monitor the course of amelioration of
CC   malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC   which are used in the exemplification of the present invention
XX
SQ   Sequence 7 AA;

  Query Match             95.2%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 20
AAM46173
ID   AAM46173 standard; peptide; 7 AA.
XX
AC   AAM46173;
XX
DT   25-OCT-2001  (first entry)
XX
DE   H11 binding site consensus conforming peptide (CCP) #2444.
XX
KW   Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW   immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW   cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW   astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW   ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   CA2290722-A1.
XX
PD   08-JUN-2001.
XX
PF   08-DEC-1999;   99CA-02290722.
XX
PR   08-DEC-1999;   99CA-02290722.
XX
PA   (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI   Kaplan HA,  Maiti PK,  Fast DG,  Herman W,  Dan MD,  Lewis KE;
PI   Entwistle JM,  Macdonald GC;
XX
DR   WPI; 2001-425937/46.
XX
PT   Composition useful for treating and diagnosing cancer, comprises stress
PT   protein-peptide complexes associated with tumor, and isolated antigen-
PT   binding fragments of an antibody that binds specifically to the complex.
XX
PS   Example 4; Page 109; 154pp; English.
XX
CC   The present invention describes a composition (I) comprising stress
CC   protein-peptide complexes (SPPC) associated with tumours that is
CC   specifically immunogenically cross-reactive with cell surface-associated
CC   SPPCs specific to target cancer (TC). Also described is an isolated
CC   antigen-binding fragment of an antibody that binds specifically to SPPCs
CC   or a population of different SPPCs consisting of immunogenic cancer cell
CC   surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC   used in vaccine production and as a tumour-specific immunogenic response
CC   inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC   subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC   oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC   ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC   vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC   or imaging cancer cells, and to monitor the course of amelioration of
CC   malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC   which are used in the exemplification of the present invention
XX
SQ   Sequence 7 AA;

  Query Match             95.2%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 21
AAM44565
ID   AAM44565 standard; peptide; 7 AA.
XX
AC   AAM44565;
XX
DT   25-OCT-2001  (first entry)
XX
DE   H11 binding site consensus conforming peptide (CCP) #836.
XX
```

KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CA2290722-A1.
XX
PD 08-JUN-2001.
XX
PF 08-DEC-1999; 99CA-02290722.
XX
PR 08-DEC-1999; 99CA-02290722.
XX
PA (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
PI Entwistle JM, Macdonald GC;
XX
DR WPI; 2001-425937/46.
XX
PT Composition useful for treating and diagnosing cancer, comprises stress
PT protein-peptide complexes associated with tumor, and isolated antigen-
PT binding fragments of an antibody that binds specifically to the complex.
XX
PS Example 4; Page 104; 154pp; English.
XX
CC The present invention describes a composition (I) comprising stress
CC protein-peptide complexes (SPPC) associated with tumours that is
CC specifically immunogenically cross-reactive with cell surface-associated
CC SPPCs specific to target cancer (TC). Also described is an isolated
CC antigen-binding fragment of an antibody that binds specifically to SPPCs
CC or a population of different SPPCs consisting of immunogenic cancer cell
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC used in vaccine production and as a tumour-specific immunogenic response
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC or imaging cancer cells, and to monitor the course of amelioration of
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC which are used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 2 RRLN 5

RESULT 22
AAM44575
ID AAM44575 standard; peptide; 7 AA.
XX
AC AAM44575;
XX
DT 25-OCT-2001 (first entry)
XX
DE H11 binding site consensus conforming peptide (CCP) #846.
XX
KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CA2290722-A1.
XX
PD 08-JUN-2001.
XX
PF 08-DEC-1999; 99CA-02290722.
XX
PR 08-DEC-1999; 99CA-02290722.
XX
PA (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
PI Entwistle JM, Macdonald GC;
XX
DR WPI; 2001-425937/46.
XX
PT Composition useful for treating and diagnosing cancer, comprises stress
PT protein-peptide complexes associated with tumor, and isolated antigen-
PT binding fragments of an antibody that binds specifically to the complex.
XX
PS Example 4; Page 104; 154pp; English.
XX
CC The present invention describes a composition (I) comprising stress
CC protein-peptide complexes (SPPC) associated with tumours that is
CC specifically immunogenically cross-reactive with cell surface-associated
CC SPPCs specific to target cancer (TC). Also described is an isolated
CC antigen-binding fragment of an antibody that binds specifically to SPPCs
CC or a population of different SPPCs consisting of immunogenic cancer cell
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC used in vaccine production and as a tumour-specific immunogenic response
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC or imaging cancer cells, and to monitor the course of amelioration of
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC which are used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 2 RRLN 5

RESULT 23
AAM43976
ID AAM43976 standard; peptide; 7 AA.
XX
AC AAM43976;
XX
DT 25-OCT-2001 (first entry)
XX
DE H11 binding site consensus conforming peptide (CCP) #247.
XX
KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CA2290722-A1.

```
XX
PD   08-JUN-2001.
XX
PF   08-DEC-1999;   99CA-02290722.
XX
PR   08-DEC-1999;   99CA-02290722.
XX
PA   (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI   Kaplan HA,  Maiti PK,  Fast DG,  Herman W,  Dan MD,  Lewis KE;
PI   Entwistle JM,  Macdonald GC;
XX
DR   WPI; 2001-425937/46.
XX
PT   Composition useful for treating and diagnosing cancer, comprises stress
PT   protein-peptide complexes associated with tumor, and isolated antigen-
PT   binding fragments of an antibody that binds specifically to the complex.
XX
PS   Example 4; Page 101; 154pp; English.
XX
CC   The present invention describes a composition (I) comprising stress
CC   protein-peptide complexes (SPPC) associated with tumours that is
CC   specifically immunogenically cross-reactive with cell surface-associated
CC   SPPCs specific to target cancer (TC). Also described is an isolated
CC   antigen-binding fragment of an antibody that binds specifically to SPPCs
CC   or a population of different SPPCs consisting of immunogenic cancer cell
CC   surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC   used in vaccine production and as a tumour-specific immunogenic response
CC   inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC   subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC   oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC   ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC   vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC   or imaging cancer cells, and to monitor the course of amelioration of
CC   malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC   which are used in the exemplification of the present invention
XX
SQ   Sequence 7 AA;

  Query Match             95.2%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 24
AAM44570
ID   AAM44570 standard; peptide; 7 AA.
XX
AC   AAM44570;
XX
DT   25-OCT-2001  (first entry)
XX
DE   H11 binding site consensus conforming peptide (CCP) #841.
XX
KW   Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW   immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW   cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW   astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW   ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   CA2290722-A1.
XX
PD   08-JUN-2001.
XX
PF   08-DEC-1999;   99CA-02290722.
XX
PR   08-DEC-1999;   99CA-02290722.
XX
PA   (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI   Kaplan HA,  Maiti PK,  Fast DG,  Herman W,  Dan MD,  Lewis KE;
PI   Entwistle JM,  Macdonald GC;
XX
DR   WPI; 2001-425937/46.
XX
PT   Composition useful for treating and diagnosing cancer, comprises stress
PT   protein-peptide complexes associated with tumor, and isolated antigen-
PT   binding fragments of an antibody that binds specifically to the complex.
XX
PS   Example 4; Page 104; 154pp; English.
XX
CC   The present invention describes a composition (I) comprising stress
CC   protein-peptide complexes (SPPC) associated with tumours that is
CC   specifically immunogenically cross-reactive with cell surface-associated
CC   SPPCs specific to target cancer (TC). Also described is an isolated
CC   antigen-binding fragment of an antibody that binds specifically to SPPCs
CC   or a population of different SPPCs consisting of immunogenic cancer cell
CC   surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC   used in vaccine production and as a tumour-specific immunogenic response
CC   inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC   subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC   oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC   ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC   vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC   or imaging cancer cells, and to monitor the course of amelioration of
CC   malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC   which are used in the exemplification of the present invention
XX
SQ   Sequence 7 AA;

  Query Match             95.2%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 25
AAM46480
ID   AAM46480 standard; peptide; 7 AA.
XX
AC   AAM46480;
XX
DT   25-OCT-2001  (first entry)
XX
DE   H11 binding site consensus conforming peptide (CCP) #2751.
XX
KW   Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW   immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW   cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW   astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW   ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   CA2290722-A1.
XX
PD   08-JUN-2001.
XX
PF   08-DEC-1999;   99CA-02290722.
XX
PR   08-DEC-1999;   99CA-02290722.
XX
PA   (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI   Kaplan HA,  Maiti PK,  Fast DG,  Herman W,  Dan MD,  Lewis KE;
```

PI Entwistle JM, Macdonald GC;
XX
DR WPI; 2001-425937/46.
XX
PT Composition useful for treating and diagnosing cancer, comprises stress
PT protein-peptide complexes associated with tumor, and isolated antigen-
PT binding fragments of an antibody that binds specifically to the complex.
XX
PS Example 4; Page 110; 154pp; English.
XX
CC The present invention describes a composition (I) comprising stress
CC protein-peptide complexes (SPPC) associated with tumours that is
CC specifically immunogenically cross-reactive with cell surface-associated
CC SPPCs specific to target cancer (TC). Also described is an isolated
CC antigen-binding fragment of an antibody that binds specifically to SPPCs
CC or a population of different SPPCs consisting of immunogenic cancer cell
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC used in vaccine production and as a tumour-specific immunogenic response
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC or imaging cancer cells, and to monitor the course of amelioration of
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC which are used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 2 RRLN 5

RESULT 26
AAM44560
ID AAM44560 standard; peptide; 7 AA.
XX
AC AAM44560;
XX
DT 25-OCT-2001 (first entry)
XX
DE H11 binding site consensus conforming peptide (CCP) #831.
XX
KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CA2290722-A1.
XX
PD 08-JUN-2001.
XX
PF 08-DEC-1999; 99CA-02290722.
XX
PR 08-DEC-1999; 99CA-02290722.
XX
PA (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
PI Entwistle JM, Macdonald GC;
XX
DR WPI; 2001-425937/46.
XX
PT Composition useful for treating and diagnosing cancer, comprises stress
PT protein-peptide complexes associated with tumor, and isolated antigen-
PT binding fragments of an antibody that binds specifically to the complex.
XX
PS Example 4; Page 104; 154pp; English.
XX
CC The present invention describes a composition (I) comprising stress
CC protein-peptide complexes (SPPC) associated with tumours that is
CC specifically immunogenically cross-reactive with cell surface-associated
CC SPPCs specific to target cancer (TC). Also described is an isolated
CC antigen-binding fragment of an antibody that binds specifically to SPPCs
CC or a population of different SPPCs consisting of immunogenic cancer cell
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC used in vaccine production and as a tumour-specific immunogenic response
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC or imaging cancer cells, and to monitor the course of amelioration of
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC which are used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 2 RRLN 5

RESULT 27
AAM45974
ID AAM45974 standard; peptide; 7 AA.
XX
AC AAM45974;
XX
DT 25-OCT-2001 (first entry)
XX
DE H11 binding site consensus conforming peptide (CCP) #2245.
XX
KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CA2290722-A1.
XX
PD 08-JUN-2001.
XX
PF 08-DEC-1999; 99CA-02290722.
XX
PR 08-DEC-1999; 99CA-02290722.
XX
PA (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
PI Entwistle JM, Macdonald GC;
XX
DR WPI; 2001-425937/46.
XX
PT Composition useful for treating and diagnosing cancer, comprises stress
PT protein-peptide complexes associated with tumor, and isolated antigen-
PT binding fragments of an antibody that binds specifically to the complex.
XX
PS Example 4; Page 109; 154pp; English.
XX

```
CC   The present invention describes a composition (I) comprising stress
CC   protein-peptide complexes (SPPC) associated with tumours that is
CC   specifically immunogenically cross-reactive with cell surface-associated
CC   SPPCs specific to target cancer (TC). Also described is an isolated
CC   antigen-binding fragment of an antibody that binds specifically to SPPCs
CC   or a population of different SPPCs consisting of immunogenic cancer cell
CC   surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC   used in vaccine production and as a tumour-specific immunogenic response
CC   inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC   subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC   oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC   ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC   vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC   or imaging cancer cells, and to monitor the course of amelioration of
CC   malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC   which are used in the exemplification of the present invention
XX
SQ   Sequence 7 AA;

  Query Match             95.2%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 28
AAM46528
ID   AAM46528 standard; peptide; 7 AA.
XX
AC   AAM46528;
XX
DT   25-OCT-2001  (first entry)
XX
DE   H11 binding site consensus conforming peptide (CCP) #2799.
XX
KW   Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW   immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW   cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW   astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW   ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   CA2290722-A1.
XX
PD   08-JUN-2001.
XX
PF   08-DEC-1999;   99CA-02290722.
XX
PR   08-DEC-1999;   99CA-02290722.
XX
PA   (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI   Kaplan HA,  Maiti PK,  Fast DG,  Herman W,  Dan MD,  Lewis KE;
PI   Entwistle JM,  Macdonald GC;
XX
DR   WPI; 2001-425937/46.
XX
PT   Composition useful for treating and diagnosing cancer, comprises stress
PT   protein-peptide complexes associated with tumor, and isolated antigen-
PT   binding fragments of an antibody that binds specifically to the complex.
XX
PS   Example 4; Page 111; 154pp; English.
XX
CC   The present invention describes a composition (I) comprising stress
CC   protein-peptide complexes (SPPC) associated with tumours that is
CC   specifically immunogenically cross-reactive with cell surface-associated
CC   SPPCs specific to target cancer (TC). Also described is an isolated
CC   antigen-binding fragment of an antibody that binds specifically to SPPCs
CC   or a population of different SPPCs consisting of immunogenic cancer cell
CC   surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC   used in vaccine production and as a tumour-specific immunogenic response
CC   inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC   subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC   oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC   ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC   vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC   or imaging cancer cells, and to monitor the course of amelioration of
CC   malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC   which are used in the exemplification of the present invention
XX
SQ   Sequence 7 AA;

  Query Match             95.2%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 29
AAM46485
ID   AAM46485 standard; peptide; 7 AA.
XX
AC   AAM46485;
XX
DT   25-OCT-2001  (first entry)
XX
DE   H11 binding site consensus conforming peptide (CCP) #2756.
XX
KW   Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW   immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW   cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW   astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW   ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   CA2290722-A1.
XX
PD   08-JUN-2001.
XX
PF   08-DEC-1999;   99CA-02290722.
XX
PR   08-DEC-1999;   99CA-02290722.
XX
PA   (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI   Kaplan HA,  Maiti PK,  Fast DG,  Herman W,  Dan MD,  Lewis KE;
PI   Entwistle JM,  Macdonald GC;
XX
DR   WPI; 2001-425937/46.
XX
PT   Composition useful for treating and diagnosing cancer, comprises stress
PT   protein-peptide complexes associated with tumor, and isolated antigen-
PT   binding fragments of an antibody that binds specifically to the complex.
XX
PS   Example 4; Page 110; 154pp; English.
XX
CC   The present invention describes a composition (I) comprising stress
CC   protein-peptide complexes (SPPC) associated with tumours that is
CC   specifically immunogenically cross-reactive with cell surface-associated
CC   SPPCs specific to target cancer (TC). Also described is an isolated
CC   antigen-binding fragment of an antibody that binds specifically to SPPCs
CC   or a population of different SPPCs consisting of immunogenic cancer cell
CC   surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC   used in vaccine production and as a tumour-specific immunogenic response
CC   inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC   subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
```

CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC or imaging cancer cells, and to monitor the course of amelioration of
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC which are used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 2 RRLN 5

RESULT 30
AAM45665
ID AAM45665 standard; peptide; 7 AA.
XX
AC AAM45665;
XX
DT 25-OCT-2001 (first entry)
XX
DE H11 binding site consensus conforming peptide (CCP) #1936.
XX
KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CA2290722-A1.
XX
PD 08-JUN-2001.
XX
PF 08-DEC-1999; 99CA-02290722.
XX
PR 08-DEC-1999; 99CA-02290722.
XX
PA (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
PI Entwistle JM, Macdonald GC;
XX
DR WPI; 2001-425937/46.
XX
PT Composition useful for treating and diagnosing cancer, comprises stress
PT protein-peptide complexes associated with tumor, and isolated antigen-
PT binding fragments of an antibody that binds specifically to the complex.
XX
PS Example 4; Page 107; 154pp; English.
XX
CC The present invention describes a composition (I) comprising stress
CC protein-peptide complexes (SPPC) associated with tumours that is
CC specifically immunogenically cross-reactive with cell surface-associated
CC SPPCs specific to target cancer (TC). Also described is an isolated
CC antigen-binding fragment of an antibody that binds specifically to SPPCs
CC or a population of different SPPCs consisting of immunogenic cancer cell
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC used in vaccine production and as a tumour-specific immunogenic response
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC or imaging cancer cells, and to monitor the course of amelioration of
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC which are used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 2 RRLN 5

RESULT 31
AAM46523
ID AAM46523 standard; peptide; 7 AA.
XX
AC AAM46523;
XX
DT 25-OCT-2001 (first entry)
XX
DE H11 binding site consensus conforming peptide (CCP) #2794.
XX
KW Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN CA2290722-A1.
XX
PD 08-JUN-2001.
XX
PF 08-DEC-1999; 99CA-02290722.
XX
PR 08-DEC-1999; 99CA-02290722.
XX
PA (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI Kaplan HA, Maiti PK, Fast DG, Herman W, Dan MD, Lewis KE;
PI Entwistle JM, Macdonald GC;
XX
DR WPI; 2001-425937/46.
XX
PT Composition useful for treating and diagnosing cancer, comprises stress
PT protein-peptide complexes associated with tumor, and isolated antigen-
PT binding fragments of an antibody that binds specifically to the complex.
XX
PS Example 4; Page 111; 154pp; English.
XX
CC The present invention describes a composition (I) comprising stress
CC protein-peptide complexes (SPPC) associated with tumours that is
CC specifically immunogenically cross-reactive with cell surface-associated
CC SPPCs specific to target cancer (TC). Also described is an isolated
CC antigen-binding fragment of an antibody that binds specifically to SPPCs
CC or a population of different SPPCs consisting of immunogenic cancer cell
CC surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC used in vaccine production and as a tumour-specific immunogenic response
CC inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC or imaging cancer cells, and to monitor the course of amelioration of
CC malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC which are used in the exemplification of the present invention
XX
SQ Sequence 7 AA;

Query Match 95.2%; Score 20; DB 4; Length 7;

```
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 32
AAM46526
ID   AAM46526 standard; peptide; 7 AA.
XX
AC   AAM46526;
XX
DT   25-OCT-2001  (first entry)
XX
DE   H11 binding site consensus conforming peptide (CCP) #2797.
XX
KW   Antigen-binding; tumour; diagnosis; stress protein-peptide complex; SPPC;
KW   immunogenically cross-reactive; cancer; immunogenic cancer cell;
KW   cytostatic; vaccine; tumour-specific immunogenic response inducer;
KW   astrocytoma; fibrosarcoma; myxosarcoma; liposarcoma; oligodendroglioma;
KW   ependymoma; medulloblastoma; primitive neural ectodermal tumour.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   CA2290722-A1.
XX
PD   08-JUN-2001.
XX
PF   08-DEC-1999;   99CA-02290722.
XX
PR   08-DEC-1999;   99CA-02290722.
XX
PA   (NOVO-) NOVOPHARM BIOTECH INC.
XX
PI   Kaplan HA,  Maiti PK,  Fast DG,  Herman W,  Dan MD,  Lewis KE;
PI   Entwistle JM,  Macdonald GC;
XX
DR   WPI; 2001-425937/46.
XX
PT   Composition useful for treating and diagnosing cancer, comprises stress
PT   protein-peptide complexes associated with tumor, and isolated antigen-
PT   binding fragments of an antibody that binds specifically to the complex.
XX
PS   Example 4; Page 111; 154pp; English.
XX
CC   The present invention describes a composition (I) comprising stress
CC   protein-peptide complexes (SPPC) associated with tumours that is
CC   specifically immunogenically cross-reactive with cell surface-associated
CC   SPPCs specific to target cancer (TC). Also described is an isolated
CC   antigen-binding fragment of an antibody that binds specifically to SPPCs
CC   or a population of different SPPCs consisting of immunogenic cancer cell
CC   surface-associated SPPC of TC. (I) has cytostatic activity and can be
CC   used in vaccine production and as a tumour-specific immunogenic response
CC   inducer. (I) is useful for treating 71 types of cancers or tumours in a
CC   subject, such as astrocytoma, fibrosarcoma, myxosarcoma, liposarcoma,
CC   oligodendroglioma, ependymoma, medulloblastoma, and primitive neural
CC   ectodermal tumour (PNET). (I) is useful as cancer immunogen including
CC   vaccines. (I) is useful for diagnostic and palliative use, for detecting
CC   or imaging cancer cells, and to monitor the course of amelioration of
CC   malignancy in an individual. AAM43707 to AAM47109 represent peptides
CC   which are used in the exemplification of the present invention
XX
SQ   Sequence 7 AA;

  Query Match             95.2%;  Score 20;  DB 4;  Length 7;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 33
AAW57007
ID   AAW57007 standard; peptide; 8 AA.
XX
AC   AAW57007;
XX
DT   28-JUL-1998  (first entry)
XX
DE   Enzyme inhibitor peptide SEQ ID NO:208.
XX
KW   Enzyme inhibitor; t-PA; u-PA; chymotrypsin; serine protease; active;
KW   latent; substrate subtraction phage display peptide library;
KW   identification; kinase; phosphatase; serpin.
XX
OS   Homo sapiens.
XX
PN   WO9747314-A1.
XX
PD   18-DEC-1997.
XX
PF   10-JUN-1997;   97WO-US009760.
XX
PR   10-JUN-1996;   96US-0019495P.
XX
PA   (SCRI ) SCRIPPS RES INST.
XX
PI   Madison EL,  Ke S;
XX
DR   WPI; 1998-062746/06.
XX
PT   Substrate subtraction phage display peptide libraries - used to
PT   distinguish between active and latent forms of enzyme, e.g. serine
PT   protease.
XX
PS   Claim 25; Page 111; 138pp; English.
XX
CC   The present sequence represents an enzyme inhibitor peptide used in the
CC   method of the invention to distinguish between t-PA and u-PA. The present
CC   invention describes a substrate subtraction library for the
CC   identification of peptide substrates selective between a first enzyme
CC   (E1) and a second enzyme (E2), comprising a collection different
CC   peptides, substantially lacking peptides that are effective substrates
CC   for E1. Also described are: (1) a method (M1) for identifying peptide
CC   substrates selective between a first enzyme (E1) and a second enzyme (E2)
CC   ; (2) a compound comprising the amino acid sequence of a peptide
CC   identified by M1; (3) a polypeptide for use as an enzyme inhibitor
CC   comprising one of 237 amino acid sequences (see AAW56801 to AAW56947, and
CC   AAW56949 to AAW57038); (4) a recombinant DNA vector comprising DNA (I)
CC   encoding a protease inhibitor including the sequence identified by the M1
CC   ; (5) a prokaryotic or eukaryotic cell containing the vector of (4); (6)
CC   an antibody (Ab) immunoreactive with at least one of the peptides
CC   identified by M1; and (7) a diagnostic assay for distinguishing between
CC   active and latent forms of protease inhibitors, that uses (Ab). The
CC   library and method are used for distinguishing between active and latent
CC   forms of enzyme inhibitors, e.g. proteases, kinases and phosphatases.
CC   (Ab) are used for affinity purification of recombinant peptides and in
CC   the identification of naturally occurring protease inhibitors. Enzyme-
CC   inhibiting peptides identified can be used to treat a serpin deficiency
CC   or a disorder of serine proteases
XX
SQ   Sequence 8 AA;

  Query Match             95.2%;  Score 20;  DB 2;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          4 RRLN 7
```

```
RESULT 34
AAU05744
ID   AAU05744 standard; protein; 8 AA.
XX
AC   AAU05744;
XX
DT   21-NOV-2001  (first entry)
XX
DE   p21 C-terminus derived peptide #103.
XX
KW   Human; p21WAFI; cyclin dependent protein kinase; CDK2; cyclin A;
KW   inhibitor; proliferative disorder; cancer; leukaemia; drug screening.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
FH   Key             Location/Qualifiers
FT   Modified-site   1
FT                   /note= "The N-terminus is hydrogenated"
FT   Modified-site   8
FT                   /label= OTHER
FT                   /note= "Other= parafluorophenylalanine. C-terminal amide"
XX
PN   WO200140142-A2.
XX
PD   07-JUN-2001.
XX
PF   29-NOV-2000; 2000WO-GB004550.
XX
PR   30-NOV-1999;   99GB-00028323.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2001-488493/53.
XX
PT   New p21 derived peptides and their variants, particularly useful as
PT   selective inhibitors of CDK2/cyclin interaction for treating
PT   proliferative disorders e.g. cancers and leukemias, and in assays for
PT   identifying CDK/cyclin inhibitors.
XX
PS   Claim 25; Page 89; 102pp; English.
XX
CC   The invention relates to peptide and their variants derived from p21WAFI,
CC   which are inhibitors of CDK2 activity by binding to G1 and S phase
CC   specific cyclins which activate CDK2; selective inhibitors of CDK2/cyclin
CC   complexes, particularly CDK2/cyclin A or E complexes. The variants of the
CC   peptide may have further amino acids at either end or have up to 7 amino
CC   acids deleted, provided the motif XLXF is retained. The peptides are
CC   specific regions of p21WAFI that bind to G1 and S phase specific cyclins,
CC   preferably cyclins which activate CDK2. One of the peptides corresponds
CC   to p21(149-159). The peptides are used for treating proliferative
CC   disorders, e.g. cancers and leukaemias. The peptides are also for
CC   identifying substances which interfere with protein-protein interactions
CC   involving cyclins (i.e. cyclin A, E or D), especially CDK/cyclin
CC   interactions, and which are capable of inhibiting CDK2 and/or CDK4
CC   activity. P21 peptides other than p21(149-159) competitively inhibit the
CC   binding of peptide p21(149-159) to cyclin and may be used to identify
CC   substances that bind to, or inhibit peptide- cyclin interactions.
CC   Substances for screening in the assays include antibody products specific
CC   for p21 or cyclin binding regions, combinatorial libraries and single
CC   compound collections. The present sequence is a peptide derived from the
CC   C-terminus of p21
XX
SQ   Sequence 8 AA;

  Query Match             95.2%;  Score 20;  DB 4;  Length 8;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          4 RRLN 7


RESULT 35
ADZ71972
ID   ADZ71972 standard; peptide; 8 AA.
XX
AC   ADZ71972;
XX
DT   14-JUL-2005  (first entry)
XX
DE   p21-derived peptide #557.
XX
KW   CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW   neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS   Synthetic.
XX
PN   WO2005040802-A2.
XX
PD   06-MAY-2005.
XX
PF   20-OCT-2004; 2004WO-GB004431.
XX
PR   20-OCT-2003; 2003GB-00024466.
PR   02-FEB-2004; 2004US-00771242.
XX
PA   (CYCL-) CYCLACEL LTD.
XX
PI   Zheleva DI,  Fischer PM,  Mcinnes C,  Andrews MJI,  Chan WC;
PI   Atkinson GE;
XX
DR   WPI; 2005-355897/36.
XX
PT   New peptide inhibitors of cyclin dependent kinases derived from the C-
PT   terminal region of p21, useful in preparing a medicament for treating a
PT   proliferative disorder such as cancer.
XX
PS   Example 25; Page 81; 112pp; English.
XX
CC   The invention relates to a peptide or its variant comprising formula: A-
CC   (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC   natural or unnatural amino acid residue having a side chain comprising at
CC   least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC   each of B or D is independently an amino acid residue selected from
CC   arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC   isoleucine or alanine; C is a natural or unnatural amino acid residue
CC   having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC   containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC   or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC   Also described are: a pharmaceutical composition comprising the peptide
CC   admixed with a diluent, an excipient or a carrier; an assay for
CC   identifying candidate substances capable of binding to a cyclin
CC   associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC   assay for identifying compounds that interact a cyclin or a cyclin when
CC   complexed with the physiologically relevant CDK; and a method of using a
CC   cyclin in a drug screening assay. The assay for identifying candidate
CC   substances capable of binding to a cyclin associated with a G1 control
CC   CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC   a peptide as defined above, the cyclin, the CDK and the candidate
CC   substance, under conditions where, in the absence of the candidate
CC   substance being an inhibitor of interaction of the cyclin/CDK
CC   interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC   any change in the expected binding of the peptide and the cyclin. The
CC   assay for identifying compounds that interact a cyclin or a cyclin when
CC   complexed with the physiologically relevant CDK comprises: incubating a
CC   candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC   and detecting binding of either the candidate compound or the peptide
CC   with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC   comprises use of a three-dimensional model of a cyclin and a candidate
CC   compound. At least one of the assay components is bound to a solid phase.
```

CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 8 AA;

Query Match 95.2%; Score 20; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 4 RRLN 7

RESULT 36
ADZ71569
ID ADZ71569 standard; peptide; 8 AA.
XX
AC ADZ71569;
XX
DT 14-JUL-2005 (first entry)
XX
DE p21-derived peptide #154.
XX
KW CDK inhibitor; Cyclin-dependent kinase-2 inhibitor; p21; cancer;
KW neoplasm; cytostatic; pharmaceutical; drug screening.
XX
OS Synthetic.
XX
PN WO2005040802-A2.
XX
PD 06-MAY-2005.
XX
PF 20-OCT-2004; 2004WO-GB004431.
XX
PR 20-OCT-2003; 2003GB-00024466.
PR 02-FEB-2004; 2004US-00771242.
XX
PA (CYCL-) CYCLACEL LTD.
XX
PI Zheleva DI, Fischer PM, Mcinnes C, Andrews MJI, Chan WC;
PI Atkinson GE;
XX
DR WPI; 2005-355897/36.
XX
PT New peptide inhibitors of cyclin dependent kinases derived from the C-
PT terminal region of p21, useful in preparing a medicament for treating a
PT proliferative disorder such as cancer.
XX
PS Disclosure; Page 14; 112pp; English.
XX
CC The invention relates to a peptide or its variant comprising formula: A-
CC (B) m -C-(D) n -E , where m or n are each independently 0 or 1; A is a
CC natural or unnatural amino acid residue having a side chain comprising at
CC least one H-bond acceptor moiety and at least one H-bond donor moiety;
CC each of B or D is independently an amino acid residue selected from
CC arginine, glycine, citrulline, glutamine, serine, lysine, asparagine,
CC isoleucine or alanine; C is a natural or unnatural amino acid residue
CC having a branched or unbranched C 1 -C 6 alkylene side chain optionally
CC containing a H-bond donor or a H-bond acceptor moiety; and E is a natural
CC or unnatural amino acid residue having an aryl or heteroaryl side chain.
CC Also described are: a pharmaceutical composition comprising the peptide
CC admixed with a diluent, an excipient or a carrier; an assay for
CC identifying candidate substances capable of binding to a cyclin
CC associated with a G1 control CDK enzyme and/or inhibiting the enzyme; an
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK; and a method of using a
CC cyclin in a drug screening assay. The assay for identifying candidate
CC substances capable of binding to a cyclin associated with a G1 control
CC CDK enzyme and/or inhibiting the enzyme comprises: bringing into contact
CC a peptide as defined above, the cyclin, the CDK and the candidate
CC substance, under conditions where, in the absence of the candidate
CC substance being an inhibitor of interaction of the cyclin/CDK
CC interaction, the peptidomimetic would bind to the cyclin; and monitoring
CC any change in the expected binding of the peptide and the cyclin. The
CC assay for identifying compounds that interact a cyclin or a cyclin when
CC complexed with the physiologically relevant CDK comprises: incubating a
CC candidate compound and the peptide and a cyclin or cyclin/CDK complex;
CC and detecting binding of either the candidate compound or the peptide
CC with the cyclin. The cyclin is cyclin A, cyclin E or cyclin D. The assay
CC comprises use of a three-dimensional model of a cyclin and a candidate
CC compound. At least one of the assay components is bound to a solid phase.
CC The peptidomimetic is labeled such as to emit a signal when bound to the
CC cyclin. The cyclin is labeled such as to emit a signal when bound to the
CC peptide. One of the assay components is labeled with a fluorescence
CC emitter and the signal is detected using fluorescence polarization
CC techniques. Using a cyclin in a drug screening assay comprises: selecting
CC a candidate compound by performing rational drug design with a three-
CC dimensional model of the cyclin, where the selecting is performed in
CC conjunction with computer modeling; contacting the candidate compound
CC with the cyclin; and detecting the binding of the candidate compound for
CC the cyclin groove. A potential drug is selected on the basis of its
CC having a greater affinity for the cyclin groove than that of the peptide.
CC The method of detection comprises monitoring G0 and/or G1/S cell cycle,
CC cell cycle-related apoptosis, suppression of E2F transcription factor,
CC hypophosphorylation of cellular pRb, or in vitro anti-proliferative
CC effects. The peptide is useful in preparing a medicament for treating a
CC proliferative disorder, e.g., cancer. The present sequence represents a
CC p21-derived peptide of the invention.
XX
SQ Sequence 8 AA;

Query Match 95.2%; Score 20; DB 9; Length 8;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 4 RRLN 7

RESULT 37
AAE18742
ID AAE18742 standard; peptide; 9 AA.
XX
AC AAE18742;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human leucocyte antigen (HLA) class I epitope #6.
XX
KW Human; lung tumour associated antigen; CASB761; vaccine; lung cancer;
KW immunotherapeutic; lung preneoplastic lesion; autoimmune disease;
KW gene therapy; cytostatic; immunosuppressive; human leucocyte antigen;
KW HLA; epitope.
XX
OS Homo sapiens.
XX
PN WO200206338-A1.
XX
PD 24-JAN-2002.

XX
PF 11-JUL-2001; 2001WO-EP007967.
XX
PR 17-JUL-2000; 2000GB-00017512.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cassart J, Gaulis S, Vinals Y De BassolsC;
XX
DR WPI; 2002-179782/23.
XX
PT Vaccine composition for treating cancer, in particular lung cancer,
PT autoimmune diseases and other related conditions, comprises a lung tumor
PT associated antigen, especially CASB761 polypeptide.
XX
PS Claim 9; Page 65; 92pp; English.
XX
CC The invention relates to vaccines comprising lung tumour associated
CC antigen referred as CASB761 and its polynucleotide. CASB761 and its DNA
CC are useful in the manufacture of a vaccine for immunotherapeutically
CC treating a patient suffering from or susceptible to lung cancer, lung
CC preneoplastic lesions or other related conditions. Vaccines of the
CC invention are useful in medicine, for treating cancer, particularly lung
CC cancer, autoimmune diseases and other related conditions. CASB761
CC polynucleotides and their proteins are useful as diagnostic reagents, to
CC diagnose different forms and states of cancer, in staging cancerous
CC disorder and grading the nature of the cancerous tissue. An antibody
CC immunospecific for CASB761 is useful to isolate and to identify clones
CC expressing CASB761 protein or to purify the polypeptide by affinity
CC chromatography and to treat or prevent, particularly lung cancer,
CC autoimmune disease and related conditions. CASB761 DNA is used in gene
CC therapy. The present sequence is human leucocyte antigen (HLA) class I
CC epitope. This sequence is used to incorporate an epitope of CASB761
CC protein
XX
SQ Sequence 9 AA;

Query Match 95.2%; Score 20; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 5 RRLN 8

RESULT 38
AAE18741
ID AAE18741 standard; peptide; 9 AA.
XX
AC AAE18741;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human leucocyte antigen (HLA) class I epitope #5.
XX
KW Human; lung tumour associated antigen; CASB761; vaccine; lung cancer;
KW immunotherapeutic; lung preneoplastic lesion; autoimmune disease;
KW gene therapy; cytostatic; immunosuppressive; human leucocyte antigen;
KW HLA; epitope.
XX
OS Homo sapiens.
XX
PN WO200206338-A1.
XX
PD 24-JAN-2002.
XX
PF 11-JUL-2001; 2001WO-EP007967.
XX
PR 17-JUL-2000; 2000GB-00017512.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cassart J, Gaulis S, Vinals Y De BassolsC;
XX
DR WPI; 2002-179782/23.
XX
PT Vaccine composition for treating cancer, in particular lung cancer,
PT autoimmune diseases and other related conditions, comprises a lung tumor
PT associated antigen, especially CASB761 polypeptide.
XX
PS Claim 9; Page 65; 92pp; English.
XX
CC The invention relates to vaccines comprising lung tumour associated
CC antigen referred as CASB761 and its polynucleotide. CASB761 and its DNA
CC are useful in the manufacture of a vaccine for immunotherapeutically
CC treating a patient suffering from or susceptible to lung cancer, lung
CC preneoplastic lesions or other related conditions. Vaccines of the
CC invention are useful in medicine, for treating cancer, particularly lung
CC cancer, autoimmune diseases and other related conditions. CASB761
CC polynucleotides and their proteins are useful as diagnostic reagents, to
CC diagnose different forms and states of cancer, in staging cancerous
CC disorder and grading the nature of the cancerous tissue. An antibody
CC immunospecific for CASB761 is useful to isolate and to identify clones
CC expressing CASB761 protein or to purify the polypeptide by affinity
CC chromatography and to treat or prevent, particularly lung cancer,
CC autoimmune disease and related conditions. CASB761 DNA is used in gene
CC therapy. The present sequence is human leucocyte antigen (HLA) class I
CC epitope. This sequence is used to incorporate an epitope of CASB761
CC protein
XX
SQ Sequence 9 AA;

Query Match 95.2%; Score 20; DB 5; Length 9;
Best Local Similarity 100.0%; Pred. No. 2e+06;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 2 RRLN 5

RESULT 39
AAE18780
ID AAE18780 standard; peptide; 9 AA.
XX
AC AAE18780;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human leucocyte antigen (HLA) class II epitope #21.
XX
KW Human; lung tumour associated antigen; CASB761; vaccine; lung cancer;
KW immunotherapeutic; lung preneoplastic lesion; autoimmune disease;
KW gene therapy; cytostatic; immunosuppressive; human leucocyte antigen;
KW HLA; epitope.
XX
OS Homo sapiens.
XX
PN WO200206338-A1.
XX
PD 24-JAN-2002.
XX
PF 11-JUL-2001; 2001WO-EP007967.
XX
PR 17-JUL-2000; 2000GB-00017512.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cassart J, Gaulis S, Vinals Y De BassolsC;
XX
DR WPI; 2002-179782/23.
XX
PT Vaccine composition for treating cancer, in particular lung cancer,
PT autoimmune diseases and other related conditions, comprises a lung tumor
PT associated antigen, especially CASB761 polypeptide.

```
XX
PS   Claim 9; Page 68; 92pp; English.
XX
CC   The invention relates to vaccines comprising lung tumour associated
CC   antigen referred as CASB761 and its polynucleotide. CASB761 and its DNA
CC   are useful in the manufacture of a vaccine for immunotherapeutically
CC   treating a patient suffering from or susceptible to lung cancer, lung
CC   preneoplastic lesions or other related conditions. Vaccines of the
CC   invention are useful in medicine, for treating cancer, particularly lung
CC   cancer, autoimmune diseases and other related conditions. CASB761
CC   polynucleotides and their proteins are useful as diagnostic reagents, to
CC   diagnose different forms and states of cancer, in staging cancerous
CC   disorder and grading the nature of the cancerous tissue. An antibody
CC   immunospecific for CASB761 is useful to isolate and to identify clones
CC   expressing CASB761 protein or to purify the polypeptide by affinity
CC   chromatography and to treat or prevent, particularly lung cancer,
CC   autoimmune disease and related conditions. CASB761 DNA is used in gene
CC   therapy. The present sequence is human leucocyte antigen (HLA) class II
CC   epitope. This sequence is used to incorporate an epitope of CASB761
CC   protein
XX
SQ   Sequence 9 AA;

  Query Match             95.2%;  Score 20;  DB 5;  Length 9;
  Best Local Similarity   100.0%;  Pred. No. 2e+06;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          6 RRLN 9


RESULT 40
AAB80739
ID   AAB80739 standard; peptide; 11 AA.
XX
AC   AAB80739;
XX
DT   02-MAY-2001  (first entry)
XX
DE   Semenogelin II hK2 cleavage site #6.
XX
KW   Cleavage; kallikrein 2; hK2; prodrug.
XX
OS   Synthetic.
XX
PN   WO200109165-A2.
XX
PD   08-FEB-2001.
XX
PF   28-JUL-2000; 2000WO-US040496.
XX
PR   29-JUL-1999;   99US-0146316P.
XX
PA   (UYJO ) UNIV JOHNS HOPKINS.
XX
PI   Denmeade SR,  Isaacs JT,  Lilja H,  Christensen SB;
XX
DR   WPI; 2001-191450/19.
XX
PT   New peptides containing cleavage sites specifically cleaved by human
PT   kallikrein 2, useful for producing prodrugs which treat hK2-producing
PT   cell proliferative disorders without exhibiting non-specific toxicity.
XX
PS   Disclosure; Fig 1; 38pp; English.
XX
CC   The present invention relates to a peptide comprising an amino acid
CC   sequence having a cleavage site specific for an enzyme having a
CC   proteolytic activity of human kallikrein 2 (hK2), and which is up to 20
CC   amino acids in length. The invention is useful for producing a prodrug
CC   which involves linking a drug which contains a primary amine to the
CC   peptide, in which the linking of the peptide to the drug inhibits the
CC   therapeutic activity of the drug
XX
SQ   Sequence 11 AA;

  Query Match             95.2%;  Score 20;  DB 4;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          7 RRLN 10


RESULT 41
ABW01200
ID   ABW01200 standard; peptide; 11 AA.
XX
AC   ABW01200;
XX
DT   15-JAN-2004  (first entry)
XX
DE   Saccharomyces cerevisiae consensus peptide #13.
XX
KW   Modulator of translation termination; MTT1; helicase B; antiviral;
KW   therapy; HCSB; nonsense mutation; yeast.
XX
OS   Saccharomyces cerevisiae.
XX
PN   US6630294-B1.
XX
PD   07-OCT-2003.
XX
PF   22-JUL-1999;   99US-00359268.
XX
PR   22-JUL-1998;   98US-0093685P.
XX
PA   (UYNE-) UNIV NEW JERSEY MEDICINE & DENTISTRY.
XX
PI   Peltz S,  Czaplinski K,  Dinman JD;
XX
DR   WPI; 2003-810549/76.
XX
PT   Identifying an agent that increases nonsense suppression, for antiviral
PT   therapy, by contacting modulator of translation termination (Mtt1) in
PT   Saccharomyces cerevisiae with a test agent, and detecting specific
PT   binding to Mtt1.
XX
PS   Disclosure; Col 49; 0pp; English.
XX
CC   The invention relates to a method of identifying an agent that increases
CC   nonsense suppression, by contacting modulator of translation termination
CC   (MTT1) also referred to as helicase B (HCSB) in Saccharomyces cerevisiae.
CC   The method is useful for identifying compositions or agents which
CC   increase nonsense suppression. The invention may also be used for
CC   antiviral therapy and for suppression of pathological nonsense mutations.
CC   The present sequence is Saccharomyces cerevisiae consensus peptide
XX
SQ   Sequence 11 AA;

  Query Match             95.2%;  Score 20;  DB 7;  Length 11;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          1 RRLN 4


RESULT 42
ADP44126
ID   ADP44126 standard; peptide; 11 AA.
XX
AC   ADP44126;
XX
```

DT 18-NOV-2004 (first entry)
XX
DE Yeast translation termination modulation protein related peptide #15.
XX
KW gene therapy; translation termination; RNA helicase; MTT1;
KW frameshift frequency; aberrant transcript degradation;
KW peptidyl transferase modulation; beta-thalassaemia; beta-globin;
KW Duchenne/Becker Muscular Dystrophy; Haemophilia A; Haemophilia B;
KW Von Willebrand Disease; Osteogenesis Imperfecta; Breast cancer;
KW Ovarian Cancer; Wilms Tumour; Hirschsprung disease; Cystic fibrosis;
KW Kidney Stone; Familial hypercholesterolaemia; Retinitis Pigmentosa;
KW Neurofibromatosis; Retinoblastoma; ATM; Costmann Disease; yeast.
XX
OS Saccharomyces cerevisiae.
XX
PN US2004115787-A1.
XX
PD 17-JUN-2004.
XX
PF 28-AUG-2003; 2003US-00652334.
XX
PR 22-JUL-1998; 98US-0093685P.
PR 22-JUL-1999; 99US-00359268.
XX
PA (PELT/) PELTZ S.
PA (CZAP/) CZAPLINSKI K.
PA (DINM/) DINMAN J D.
XX
PI Peltz S, Czaplinski K, Dinman JD;
XX
DR WPI; 2004-449400/42.
XX
PT Identifying a test composition or agent that modulates the efficiency of
PT translation termination comprises contacting the MTT1 with the test
PT composition or agent, and determining if the test composition or agent
PT inhibits the MTT1.
XX
PS Disclosure; SEQ ID NO 24; 41pp; English.
XX
CC The invention relates to a method of identifying a test composition that
CC modulates the efficiency of translation termination comprising contacting
CC the RNA helicase MTT1 with a composition or agent under conditions
CC permitting binding between the MTT1 and the composition, detecting
CC specific binding of the test composition or agent to the MTT1, and
CC determining if the test composition or agent inhibits the MTT1. The
CC composition and methods are useful for modulating the fidelity of
CC translation termination or for identifying agents that: affect the
CC functional activity of mRNAs by altering frameshift frequency, permit
CC monitoring of a termination event, promote degradation of aberrant
CC transcripts, and provide modulators (inhibitors/stimulators) of peptidyl
CC transferase activity during initiation, elongation, termination and mRNA
CC degradation of translation. The agents, which may be antagonists or
CC agonists, are useful in screening, diagnostic and therapeutic purposes,
CC for diseases or conditions resulting from or cause premature translation,
CC such as beta-thalassaemia, beta-globin, Duchenne/Becker Muscular
CC Dystrophy, Haemophilia A, Haemophilia B, Von Willebrand Disease,
CC Osteogenesis Imperfecta, Breast cancer, Ovarian Cancer, Wilms Tumour,
CC Hirschsprung disease, Cystic fibrosis, Kidney Stones, Familial
CC hypercholesterolaemia, Retinitis Pigmentosa, or Neurofibromatosis,
CC Retinoblastoma, ATM or Costmann Disease. The present sequence represents
CC the amino acid sequence of a yeast translation termination modulation
CC protein related peptide. Note: this sequence appears in the sequence
CC listing but no reference is made to it in the main body of the
CC specification.
XX
SQ Sequence 11 AA;

Query Match 95.2%; Score 20; DB 8; Length 11;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 43
AAB59999
ID AAB59999 standard; peptide; 12 AA.
XX
AC AAB59999;
XX
DT 05-NOV-2001 (first entry)
XX
DE Internalising peptide SEQ ID NO: 14.
XX
KW Internalising peptide; transport; apoptosis; arthritis; cancer;
KW stem cell; cell differentiation; immune response stimulation;
KW HIV vaccine.
XX
OS Synthetic.
XX
PN WO200115511-A2.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-US024034.
XX
PR 01-SEP-1999; 99US-0151980P.
PR 13-MAR-2000; 2000US-0188944P.
XX
PA (UYPI-) UNIV PITTSBURGH.
XX
PI Robbins PD, Mi Z, Frizzell R, Glorioso JC, Gambotto A;
XX
DR WPI; 2001-273309/28.
XX
PT Peptides that facilitate uptake and cytoplasmic and/or nuclear transport
PT of proteins, DNA and viruses, useful, e.g. for facilitating uptake of
PT antigens in immunogenic compositions.
XX
PS Claim 1; Page 116; 129pp; English.
XX
CC The present invention provides the sequences of 75 peptides which
CC facilitate the uptake and transport of viruses, proteins and nucleic
CC acids. These internalising peptides can be used for transport into the
CC cytoplasm or the nucleus. They are useful for facilitating uptake into
CC the cell, inducing apoptosis, for example in the treatment of arthritis
CC and cancer, to expand a population of stem cells or differentiated cells,
CC to stimulate cell differentiation, facilitate the integration of AAV into
CC the genome of a cell, and to stimulate an immune response, for example in
CC the case of a HIV vaccine. The present sequence is one of the peptides of
CC the invention
XX
SQ Sequence 12 AA;

Query Match 95.2%; Score 20; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 8 RRLN 11

RESULT 44
ADA88834
ID ADA88834 standard; peptide; 12 AA.
XX
AC ADA88834;
XX
DT 20-NOV-2003 (first entry)
XX
DE Internalised peptide pep14 SEQ ID NO:14.
XX
KW internalising peptide; cytostatic; antiinflammatory; immunomodulator;

KW antiarthritic; cytoplasmic transport; nuclear transport;
KW peptide-cargo complex; apoptosis; arthritic; tumour; differentiation;
KW immune response; vaccine; inflammation; necrosis; transplantation;
KW cystic fibrosis; lung inflammation; gene therapy.
XX
OS Synthetic.
XX
PN WO2003068942-A2.
XX
PD 21-AUG-2003.
XX
PF 12-FEB-2003; 2003WO-US004632.
XX
PR 13-FEB-2002; 2002US-00075869.
XX
PA (UYPI-) UNIV PITTSBURGH.
XX
PI Robbins PD, Mi Z, Frizzel R, Glorioso JC, Gambotto A, Mai JC;
XX
DR WPI; 2003-697526/66.
XX
PT New internalizing peptides, useful for facilitating the delivery, uptake
PT and cytoplasmic and/or nuclear transport of proteins, DNA or viruses into
PT a target cell, for inducing apoptosis in arthritic or tumor cells, or in
PT gene therapy.
XX
PS Example 3; Page 18; 171pp; English.
XX
CC The present invention describes an internalising peptide (I) comprising
CC any one of 14 fully defined amino acid sequences (designated P1-P14, see
CC ADA88896 to ADA88906, and ADA88917 to ADA88919). (I) has cytostatic,
CC antiinflammatory, immunomodulator and antiarthritic activities. The
CC internalising peptides are useful for facilitating the delivery, uptake
CC and cytoplasmic and/or nuclear transport of cargo, e.g. proteins, DNA or
CC viruses, into a target cell. The internalising peptides and peptide-cargo
CC complexes from the present invention are also useful for inducing
CC apoptosis in cells (e.g. arthritic cells or tumour cells), expanding a
CC population of stem cell or differentiated cells, stimulating the
CC differentiation of a population of stem cells, facilitating the
CC integration of adeno-associated virus DNA into the genome of a cell,
CC stimulating or eliciting an immune response in a subject, facilitating
CC the delivery of immunogens (e.g. vaccines), inhibiting the inflammatory
CC process, protecting tissue from apoptosis or necrosis during tissue
CC isolation prior to transplantation, facilitating transfer of proteins and
CC peptides to the lung for the treatment of cystic fibrosis or lung
CC inflammation, or in gene therapy. The present sequence represents a
CC peptide used in the exemplification of the present invention.
XX
SQ Sequence 12 AA;

```
  Query Match             95.2%;  Score 20;  DB 7;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          8 RRLN 11
```

RESULT 45
ADQ15921
ID ADQ15921 standard; peptide; 12 AA.
XX
AC ADQ15921;
XX
DT 07-OCT-2004 (first entry)
XX
DE Human Mash-1 splice variant hMash-1-deltaN76-specific epitope.
XX
KW Human; transcriptional modulator; splice variant; tumour-specific;
KW tumour-enriched; differential expression; expression pattern; cancer;
KW diagnosis; Mash-1; variant-specific antibody; autoantibody;
KW diagnostic array; prognosis; lung cancer; gastrointestinal cancer;
KW breast cancer; prostate cancer; skin cancer; sarcoma; endocrine cancer;
KW neural cancer; bladder cancer; cervical cancer; renal cancer;
KW haematopoietic cancer; cytostatic; transcription modulator inhibitor;
KW gene therapy; hMash-1-deltaN76; epitope.
XX
OS Homo sapiens.
XX
PN WO2004060302-A2.
XX
PD 22-JUL-2004.
XX
PF 24-DEC-2003; 2003WO-US041253.
XX
PR 26-DEC-2002; 2002US-0436693P.
XX
PA (CEMI-) CEMINES LLC.
XX
PI Neuman T, Palm K;
XX
DR WPI; 2004-534307/51.
XX
PT Transcription modulator splice variants, useful in diagnosing or treating
PT cancers such as lung, gastrointestinal, breast, prostate, skin,
PT endocrine, cervical and renal cancer.
XX
PS Claim 53; Page 50; 59pp; English.
XX
CC The invention relates to tumour-specific or tumour-enriched splice
CC variants of the human transcription modulators neuralised-1 (hNeu delta
CC NHR1), NeuroD1 (hNeuroD1), Irx-2 (Irx-2A) and Mash-1 (hMash-1-deltaN76)
CC and nucleic acids encoding these splice variants. The invention also
CC relates to a method of diagnosing cancer by determining the expression of
CC at least one splice variant of each of a plurality of transcription
CC modulators, or a plurality of splice variants of at least one
CC transcription modulator, where the splice variants are differentially
CC expressed compared to the wild-type isoform of the corresponding
CC transcription modulator, and where the expression pattern of the splice
CC variants is indicative of cancer. In this method, the transcription
CC modulators whose splice variant expression is analysed are selected from
CC NRSF, MDM2 TSG, RREB, ZNF207, TTF-1, GTFIIIA, HES-6, HRY, Msx2, Neu,
CC NeuroD1, Mash-1 and Irx-2. The invention further relates to a method for
CC the treatment of cancer by administration of a bioactive agent capable of
CC inhibiting the activity of one or more of the transcriptional modulator
CC splice variants; an antibody which binds to a transcription modulator
CC splice variant but not to the wild-type isoform; a peptide which binds to
CC the splice variant-specific antibody; and a diagnostic array for the
CC diagnosis of cancer comprising peptides which are capable of binding to
CC splice variant-specific autoantibodies but not to wild-type isoform-
CC specific autoantibodies. The methods and compositions of the present
CC invention are useful for the diagnosis, prognosis and/or treatment of
CC cancer, such as lung cancer, gastrointestinal cancer, breast cancer,
CC prostate cancer, skin cancer, sarcoma, endocrine cancer, neural cancer,
CC bladder cancer, cervical cancer, renal cancer and haematopoietic cancer.
CC The present sequence represents a specifically claimed peptide which
CC binds to an antibody specific for the human tumour-specific/enriched Mash
CC -1 splice variant hMash-1-deltaN76.
XX
SQ Sequence 12 AA;

```
  Query Match             95.2%;  Score 20;  DB 8;  Length 12;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          5 RRLN 8
```

RESULT 46
AAR47006
ID AAR47006 standard; protein; 14 AA.
XX
AC AAR47006;

```
XX
DT   25-MAR-2003  (revised)
DT   16-SEP-1994  (first entry)
XX
DE   LAR protein position 1302-1316.
XX
KW   Naturally-occuring; immunomodulatory protein; human; therapy; class I;
KW   major histocompatibility complex; class II; allotype; type I diabetes;
KW   autoimmune disease; rheumatoid arthritis; T-cell-mediated response;
KW   mutiple sclerosis; transplant rejection; vaccine; MHC.
XX
OS   Homo sapiens.
XX
PN   WO9404171-A1.
XX
PD   03-MAR-1994.
XX
PF   11-AUG-1993;   93WO-US007545.
XX
PR   11-AUG-1992;   92US-00925460.
PR   15-JUN-1993;   93US-00077255.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Urban RG,  Chicz RM,  Vignali DA,  Hedley ML,  Stern LJ;
PI   Strominger JL;
XX
DR   WPI; 1994-082825/10.
XX
PT   Novel immunomodulatory peptide(s) and nucleic acids - useful for
PT   treatment of auto:immune diseases, transplant rejection and for
PT   vaccination.
XX
PS   Disclosure; Page 48; 139pp; English.
XX
CC   The sequences given in AAR49291-505 and AAR46981-7038 represent peptide
CC   fragments of naturally-occuring immunomodulatory proteins. These
CC   fragments are between 10-30 residues in length and bind to a human major
CC   histocompatibility complex (MHC) class II allotype. These peptides may be
CC   used for therapy of autoimmune diseases, such as type I diabetes,
CC   rheumatoid arthritis and mutiple sclerosis, and to reduce transplant
CC   rejection. They may also be used for vaccination providing an exclusively
CC   T-cell-mediated response, which can be class I or class-II based, or
CC   both, depending on the length and character of the immunogenic peptides.
CC   (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25-MAR-2003 to
CC   correct PR field.)
XX
SQ   Sequence 14 AA;

  Query Match             95.2%;  Score 20;  DB 2;  Length 14;
  Best Local Similarity   100.0%;  Pred. No. 3.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          6 RRLN 9


RESULT 47
AAW42267
ID   AAW42267 standard; peptide; 15 AA.
XX
AC   AAW42267;
XX
DT   08-APR-1998  (first entry)
XX
DE   Biotinylated cross-linked interleukin-8 15-mer peptide ligand 7.
XX
KW   Bacteriophage peptide library; peptide epitope; therapeutic target;
KW   variegated compound library; interleukin-8; IL-8.
XX
OS   Synthetic.
XX
PN   WO9735194-A2.
XX
PD   25-SEP-1997.
XX
PF   21-MAR-1997;   97WO-US004176.
XX
PR   21-MAR-1996;   96US-00622338.
XX
PA   (HARD ) HARVARD COLLEGE.
XX
PI   Forster AC;
XX
DR   WPI; 1997-480355/44.
XX
PT   Identifying compounds which interact with target molecules - using
PT   enantiomers of the target molecules and testing of enantiomers of
PT   selected compounds.
XX
PS   Disclosure; Fig 6; 89pp; English.
XX
CC   15-mer peptides AAW42261-77 are identified as ligands of a biotinylated,
CC   cross-linked interleukin-8 (IL-8) target, using the method of the
CC   invention. This novel method identifies compounds which interact with a
CC   target molecule, and comprises contacting a screening molecule with a
CC   variegated compound library, where the screening molecule comprises solid
CC   target molecule, or the enantiomer if the target molecule is chiral.
CC   Compounds which have a desired interaction with the target molecule are
CC   selected, and the ability of their enantiomer to interact with the target
CC   molecule is tested. Ligands for a target protein can be identified by
CC   combining a D-enantiomer of a target protein (a D-target protein), and a
CC   variegated compound library, and then selecting one or more compounds
CC   from the library which have a desired binding interaction with the D-
CC   target protein. The methods can be used for identifying agonists or
CC   antagonists of targets such as receptors, enzymes, DNA binding proteins
CC   or signal transduction proteins. The methods can provide a structurally
CC   selective approach in addition to scoring for interaction of functional
CC   groups. They provide a powerful selection method that allows for the
CC   production of ligands with the same diversity as peptides but with the
CC   greatly improved pharmacokinetic profiles needed for drug activity
XX
SQ   Sequence 15 AA;

  Query Match             95.2%;  Score 20;  DB 2;  Length 15;
  Best Local Similarity   100.0%;  Pred. No. 4e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          5 RRLN 8


RESULT 48
AAW15509
ID   AAW15509 standard; peptide; 15 AA.
XX
AC   AAW15509;
XX
DT   04-JUN-1997  (first entry)
XX
DE   Peptide MIPP4.
XX
KW   Phosphorylated PS peptide; antibody; phosphoenzyme.
XX
OS   Synthetic.
XX
PN   JP09068527-A.
XX
PD   11-MAR-1997.
XX
PF   31-AUG-1995;   95JP-00224179.
XX
PR   31-AUG-1995;   95JP-00224179.
XX
```

PA (IGAK-) IGAKU SEIBUTSUGAKU KENKYUSHO KK.
XX
DR WPI; 1997-222835/20.
XX
PT Monoclonal antibody against phosphated PS peptide - used in an
PT immunoassay for determining protein phospho:enzyme activity.
XX
PS Example 4; Page 10; 13pp; Japanese.
XX
CC This sequence represents the MIPP4 peptide. This sequence is used in an
CC immunoassay method for determining protein phosphoenzyme activity. The
CC immunoassay method also uses a monoclonal antibody specific for a
CC phosphorylated PS peptide (see AAW15508). This method can determine the
CC protein phosphoenzyme activity in a sample with high specificity
XX
SQ Sequence 15 AA;

Query Match 95.2%; Score 20; DB 2; Length 15;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 2 RRLN 5

RESULT 49
ABB76940
ID ABB76940 standard; peptide; 15 AA.
XX
AC ABB76940;
XX
DT 22-JUL-2002 (first entry)
XX
DE Rat VG51-2 peptide.
XX
KW Rat; antiasthmatic; anxiolytic; antiepileptic; antihypertensive;
KW psychotropic; glutamate transporter; transporter; GABA;
KW gamma-aminobutyric acid transporter; GABA transporter; neurotransmitter;
KW asthma; anxiety; epilepsy; hypertension; psychiatric disorder;
KW neurotic disorder; VG51.
XX
OS Rattus sp.
XX
PN WO200071709-A1.
XX
PD 30-NOV-2000.
XX
PF 19-MAY-2000; 2000WO-FR001383.
XX
PR 21-MAY-1999; 99FR-00006525.
XX
PA (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX
PI Giros B, Gasnier B, Sagne C, El Mestikawy S, Hamon M;
XX
DR WPI; 2001-025160/03.
XX
PT New mammalian amino acid transporter, used e.g. to screen for
PT psychotropic agents, is high capacity but low affinity transporter of
PT gamma-aminobutyric acid.
XX
PS Claim 3; Page 44; 103pp; French.
XX
CC The present sequence is a peptide of rat VG51, a glutamate/ gamma-
CC aminobutyric acid (GABA) transporter. GABA and glutamate are
CC neurotransmitters. The transporter can be used to produce specific
CC antibodies, to screen for binding agents. Modulators of the transporter
CC are useful for treating disorders associated with deregulated
CC glutamate/GABA transport, e.g. asthma, anxiety, epilepsy, hypertension
CC and other psychiatric and neurotic disorders, while determining levels of
CC the transporter and its coding sequence can be used for diagnosis of such
CC disorders
XX
SQ Sequence 15 AA;

Query Match 95.2%; Score 20; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 1 RRLN 4

RESULT 50
AAE18774
ID AAE18774 standard; peptide; 15 AA.
XX
AC AAE18774;
XX
DT 17-MAY-2002 (first entry)
XX
DE Human leucocyte antigen (HLA) class II epitope #15.
XX
KW Human; lung tumour associated antigen; CASB761; vaccine; lung cancer;
KW immunotherapeutic; lung preneoplastic lesion; autoimmune disease;
KW gene therapy; cytostatic; immunosuppressive; human leucocyte antigen;
KW HLA; epitope.
XX
OS Homo sapiens.
XX
PN WO200206338-A1.
XX
PD 24-JAN-2002.
XX
PF 11-JUL-2001; 2001WO-EP007967.
XX
PR 17-JUL-2000; 2000GB-00017512.
XX
PA (SMIK) SMITHKLINE BEECHAM BIOLOGICALS.
XX
PI Cassart J, Gaulis S, Vinals Y De BassolsC;
XX
DR WPI; 2002-179782/23.
XX
PT Vaccine composition for treating cancer, in particular lung cancer,
PT autoimmune diseases and other related conditions, comprises a lung tumor
PT associated antigen, especially CASB761 polypeptide.
XX
PS Claim 9; Page 67; 92pp; English.
XX
CC The invention relates to vaccines comprising lung tumour associated
CC antigen referred as CASB761 and its polynucleotide. CASB761 and its DNA
CC are useful in the manufacture of a vaccine for immunotherapeutically
CC treating a patient suffering from or susceptible to lung cancer, lung
CC preneoplastic lesions or other related conditions. Vaccines of the
CC invention are useful in medicine, for treating cancer, particularly lung
CC cancer, autoimmune diseases and other related conditions. CASB761
CC polynucleotides and their proteins are useful as diagnostic reagents, to
CC diagnose different forms and states of cancer, in staging cancerous
CC disorder and grading the nature of the cancerous tissue. An antibody
CC immunospecific for CASB761 is useful to isolate and to identify clones
CC expressing CASB761 protein or to purify the polypeptide by affinity
CC chromatography and to treat or prevent, particularly lung cancer,
CC autoimmune disease and related conditions. CASB761 DNA is used in gene
CC therapy. The present sequence is human leucocyte antigen (HLA) class II
CC epitope. This sequence is used to incorporate an epitope of CASB761
CC protein
XX
SQ Sequence 15 AA;

Query Match 95.2%; Score 20; DB 5; Length 15;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 RRLN 4
              ||||
Db          8 RRLN 11
```

Search completed: January 25, 2006, 18:39:30
Job time : 129.5 secs

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:34:13 ; Search time 12.5 Seconds
(without alignments)
38.487 Million cell updates/sec

Title: US-10-771-242-295
Perfect score: 21
Sequence: 1 RRLNX 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : PIR_80:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	95.2	51	2	AF0749	hypothetical prote
2	20	95.2	51	2	S00016	protamine Z1 - sma
3	20	95.2	51	2	D98179	hypothetical prote
4	20	95.2	61	2	B84091	hypothetical prote
5	20	95.2	71	2	B64479	hypothetical prote
6	20	95.2	72	2	I39470	apolipoprotein B -
7	20	95.2	73	2	B90729	hypothetical prote
8	20	95.2	73	2	B90775	C4-type zinc finge
9	20	95.2	73	2	F91004	probable C4-type z
10	20	95.2	74	2	E82829	hypothetical prote
11	20	95.2	76	2	C64871	hypothetical prote
12	20	95.2	80	2	C91287	hypothetical prote
13	20	95.2	85	2	T08595	cysteine proteinas
14	20	95.2	87	2	S00180	spermatid protein
15	20	95.2	87	2	B37475	probable structura
16	20	95.2	89	2	C90392	conserved hypothet
17	20	95.2	90	2	H83003	conserved hypothet
18	20	95.2	91	2	A95415	hypothetical prote
19	20	95.2	92	2	AE0430	hypothetical prote
20	20	95.2	94	1	RBBP22	abc1 protein - pha
21	20	95.2	95	2	AD0229	hypothetical prote
22	20	95.2	96	2	B26074	cysteine proteinas
23	20	95.2	96	2	S51929	homeotic protein C
24	20	95.2	96	2	S51928	homeotic protein C
25	20	95.2	98	2	S62346	L71-5 protein - fr
26	20	95.2	98	2	A39437	exopolysaccharide
27	20	95.2	98	2	F95975	posttranscription
28	20	95.2	98	2	G96029	hypothetical prote
29	20	95.2	100	2	S62333	L71-1 protein - fr
30	20	95.2	102	2	T50981	hypothetical prote
31	20	95.2	103	2	E82425	hypothetical prote
32	20	95.2	103	2	S43813	hypothetical prote
33	20	95.2	105	2	T06820	lipid transfer pro
34	20	95.2	108	2	F84479	En/Spm-like transp
35	20	95.2	109	1	CABO	cytochrome-c oxida
36	20	95.2	110	2	AF0301	conserved hypothet
37	20	95.2	112	2	F71149	hypothetical prote
38	20	95.2	113	2	E83073	hypothetical prote
39	20	95.2	113	2	AD2868	hypothetical prote
40	20	95.2	114	2	AI0294	probable membrane
41	20	95.2	115	2	AF0052	hypothetical prote
42	20	95.2	115	2	S64490	hypothetical prote
43	20	95.2	116	2	T18073	hypothetical prote
44	20	95.2	118	2	H97644	hypothetical prote
45	20	95.2	121	2	B72663	hypothetical prote
46	20	95.2	121	2	G70074	hypothetical prote
47	20	95.2	123	2	G01477	ribosomal protein
48	20	95.2	123	2	B95328	hypothetical prote
49	20	95.2	124	2	T37118	probable transposa
50	20	95.2	125	2	AF2757	conserved hypothet
51	20	95.2	125	2	C97538	hypothetical prote
52	20	95.2	126	2	S47103	urease (EC 3.5.1.5
53	20	95.2	128	2	A72533	hypothetical prote
54	20	95.2	128	2	H65171	hypothetical trans
55	20	95.2	128	2	C84708	probable MYB famil
56	20	95.2	129	2	A43952	neuropeptides prec
57	20	95.2	129	2	B71068	hypothetical prote
58	20	95.2	130	2	A83400	hypothetical prote
59	20	95.2	131	2	AF0510	transcription acti
60	20	95.2	134	2	S19144	spbA protein - Bac
61	20	95.2	134	2	S77785	hypothetical prote
62	20	95.2	135	2	S75912	hypothetical prote
63	20	95.2	136	2	H71668	ribosomal protein
64	20	95.2	138	2	F81898	probable membrane
65	20	95.2	139	1	RKMLS	ribulose-bisphosph
66	20	95.2	139	2	D90303	hypothetical prote
67	20	95.2	140	2	T09267	probable tail comp
68	20	95.2	140	2	G85746	unknown protein en
69	20	95.2	140	2	D90872	hypothetical prote
70	20	95.2	141	2	T09198	probable tail comp
71	20	95.2	144	2	B83596	conserved hypothet
72	20	95.2	144	2	F28585	protein D - Escher
73	20	95.2	145	1	S05495	cytochrome-c oxida
74	20	95.2	145	1	R3HU15	ribosomal protein
75	20	95.2	145	1	R3HY15	ribosomal protein
76	20	95.2	145	1	R3RT15	ribosomal protein
77	20	95.2	145	2	B34823	ribosomal protein
78	20	95.2	145	2	C34823	ribosomal protein
79	20	95.2	145	2	A34823	ribosomal protein
80	20	95.2	146	1	S04592	cytochrome-c oxida
81	20	95.2	146	2	A95254	adc operon repress
82	20	95.2	146	2	E84608	hypothetical prote
83	20	95.2	146	2	T46753	repressor protein
84	20	95.2	147	2	G98091	conserved hypothet
85	20	95.2	147	2	C95227	conserved hypothet
86	20	95.2	147	2	T49502	hypothetical prote
87	20	95.2	148	2	S48980	hypothetical prote
88	20	95.2	149	2	F97312	hypothetical prote
89	20	95.2	150	1	OTHU5A	cytochrome-c oxida
90	20	95.2	150	2	F81280	hypothetical prote
91	20	95.2	151	2	G70377	conserved hypothet
92	20	95.2	151	2	A81094	conserved hypothet
93	20	95.2	151	2	B81847	conserved hypothet
94	20	95.2	152	2	A70800	hypothetical prote
95	20	95.2	156	2	T17274	hypothetical prote
96	20	95.2	160	2	A71300	conserved hypothet
97	20	95.2	161	2	S30937	allophycocyanin al
98	20	95.2	161	2	A30938	myosin regulatory
99	20	95.2	161	2	E71866	hypothetical prote
100	20	95.2	162	2	E83118	probable phenazine
101	20	95.2	163	2	T22481	hypothetical prote
102	20	95.2	164	2	A82393	hypothetical prote

103	20	95.2	165	2	AB1760	acetyltransferase
104	20	95.2	166	2	G98118	hypothetical prote
105	20	95.2	167	2	AB2180	DnaJ protein [impo
106	20	95.2	168	2	I40352	single-stranded DN
107	20	95.2	168	2	AB3362	single-strand bind
108	20	95.2	168	2	B83473	hypothetical prote
109	20	95.2	169	2	AH1032	hypothetical prote
110	20	95.2	170	2	D71141	hypothetical prote
111	20	95.2	170	2	T25399	hypothetical prote
112	20	95.2	170	2	T34588	hypothetical prote
113	20	95.2	171	2	H75010	hypothetical prote
114	20	95.2	171	2	H87618	hypothetical prote
115	20	95.2	172	2	A72231	dihydrofolate redu
116	20	95.2	174	2	AD2665	conserved hypothet
117	20	95.2	175	2	T39198	hypothetical prote
118	20	95.2	176	2	F71064	micrococcal nuclea
119	20	95.2	176	2	B72698	hypothetical prote
120	20	95.2	177	2	G64499	hypothetical prote
121	20	95.2	177	2	T36271	probable RNA polym
122	20	95.2	179	2	S76208	CDPdiacylglycerol-
123	20	95.2	179	2	AI2104	filament integrity
124	20	95.2	181	2	S77590	cold shock protein
125	20	95.2	184	2	S41451	regA protein - Rho
126	20	95.2	184	2	A42219	light-harvesting a
127	20	95.2	184	2	A36862	photosynthetic res
128	20	95.2	185	2	D83658	stage V sporulatio
129	20	95.2	185	2	A70075	hypothetical prote
130	20	95.2	185	2	G70425	hypothetical prote
131	20	95.2	185	2	F87279	response regulator
132	20	95.2	187	2	S32968	probable membrane
133	20	95.2	187	2	AE3478	acid tolerance reg
134	20	95.2	189	2	G70485	hypothetical prote
135	20	95.2	189	2	B69177	hypothetical prote
136	20	95.2	191	1	VUQB1B	BabR protein 1 - B
137	20	95.2	192	2	S49429	hypothetical prote
138	20	95.2	195	2	B81066	hypothetical prote
139	20	95.2	195	2	AC2583	two component resp
140	20	95.2	195	2	A97365	actR protein (AF22
141	20	95.2	196	2	G75353	hypothetical prote
142	20	95.2	196	2	F70428	probable CDP-alcoh
143	20	95.2	196	2	I50507	achaete-scute homo
144	20	95.2	197	1	YTBST	tunicamycin resist
145	20	95.2	198	2	C97447	hypothetical prote
146	20	95.2	198	2	T25436	hypothetical prote
147	20	95.2	199	2	A56548	pro-neural achaete
148	20	95.2	200	1	VUQB2B	BabR protein 2 - B
149	20	95.2	200	2	B83558	hypothetical prote
150	20	95.2	200	2	AB2123	hypothetical prote
151	20	95.2	201	2	A97327	probable membrane
152	20	95.2	202	2	S39198	nitrogenase (EC 1.
153	20	95.2	202	2	S39197	nitrogenase (EC 1.
154	20	95.2	202	2	C83859	hypothetical prote
155	20	95.2	203	2	F90302	conserved hypothet
156	20	95.2	204	2	AC1961	ATP-dependent Clp
157	20	95.2	205	2	F81718	guanylate kinase T
158	20	95.2	205	2	A71567	probable GMP kinas
159	20	95.2	205	2	F86932	conserved hypothet
160	20	95.2	206	2	C88288	protein ZK970.2 [i
161	20	95.2	206	2	A83318	hypothetical prote
162	20	95.2	207	2	AD3229	protein 6b [import
163	20	95.2	207	2	B30832	hypothetical prote
164	20	95.2	208	2	S59772	hypothetical prote
165	20	95.2	209	2	G96969	probable methyltra
166	20	95.2	209	2	C83079	hypothetical prote
167	20	95.2	210	2	F87492	ATP-dependent Clp
168	20	95.2	210	2	E97512	clpp (AF218420) [i
169	20	95.2	210	2	AE2731	ATP-dependent Clp
170	20	95.2	210	2	A95855	probable transcrip
171	20	95.2	210	2	A87022	conserved hypothet
172	20	95.2	212	1	VUQB3B	BabR protein 3 - B
173	20	95.2	212	2	T26887	hypothetical prote
174	20	95.2	213	2	T27841	hypothetical prote
175	20	95.2	213	2	AH2636	conserved hypothet
176	20	95.2	214	2	C42327	alpha 2,6-sialyltr
177	20	95.2	217	2	S65830	alpha fucosidase p
178	20	95.2	217	2	S49578	trypsin inhibitor
179	20	95.2	217	2	JQ1294	hypothetical 23.7K
180	20	95.2	218	2	T27954	hypothetical prote
181	20	95.2	218	2	S64086	hypothetical prote
182	20	95.2	219	2	A96813	hypothetical prote
183	20	95.2	219	2	I51382	achaete-scute homo
184	20	95.2	220	2	S31152	hypothetical prote
185	20	95.2	220	2	D64316	restriction modifi
186	20	95.2	221	2	T33414	hypothetical prote
187	20	95.2	222	2	T27842	hypothetical prote
188	20	95.2	222	2	T20534	hypothetical prote
189	20	95.2	224	2	E72049	conserved hypothet
190	20	95.2	224	2	F86575	CT691 hypothetical
191	20	95.2	224	2	G90448	hypothetical prote
192	20	95.2	225	2	B97580	hypothetical prote
193	20	95.2	225	2	AH2800	conserved hypothet
194	20	95.2	225	2	C70500	probable transcrip
195	20	95.2	226	2	JC5327	adhesin complex 25
196	20	95.2	227	2	AD3529	two component resp
197	20	95.2	227	2	T32894	hypothetical prote
198	20	95.2	228	2	B89889	conserved hypothet
199	20	95.2	228	2	S76711	hypothetical prote
200	20	95.2	228	2	T51147	hypothetical prote
201	20	95.2	228	2	AF3179	agrobacterium viru
202	20	95.2	229	1	G64371	conserved hypothet
203	20	95.2	229	2	H82043	guanylate kinase V
204	20	95.2	229	2	D90668	probable xanthine
205	20	95.2	229	2	F64754	probable oxidoredu
206	20	95.2	229	2	G85518	glucuronolactone r
207	20	95.2	229	2	T05150	hypothetical prote
208	20	95.2	230	2	C75412	probable cytochrom
209	20	95.2	231	2	S28186	achaete-scute locu
210	20	95.2	233	2	T51160	hypothetical prote
211	20	95.2	233	2	S11563	probable MASH-2 pr
212	20	95.2	234	2	H95253	zinc ABC transport
213	20	95.2	234	2	F98118	hypothetical prote
214	20	95.2	234	2	T46754	AdcC protein [impo
215	20	95.2	235	2	S70132	probable membrane
216	20	95.2	235	2	AG2170	hypothetical prote
217	20	95.2	237	2	G97418	hypothetical prote
218	20	95.2	238	2	AC2913	transcription regu
219	20	95.2	238	2	F97687	hypothetical prote
220	20	95.2	238	2	D86917	conserved hypothet
221	20	95.2	238	2	A48279	achaete scute prot
222	20	95.2	239	2	S07406	thaumatin homolog
223	20	95.2	239	2	B87280	conserved hypothet
224	20	95.2	239	2	AF0555	probable regulator
225	20	95.2	239	2	T46948	probable repressor
226	20	95.2	240	2	E83983	tetrahydrodipicoli
227	20	95.2	240	2	A95219	hypothetical prote
228	20	95.2	240	2	G98082	hypothetical prote
229	20	95.2	240	2	D90337	hypothetical prote
230	20	95.2	240	2	E84701	hypothetical prote
231	20	95.2	241	2	AC2083	phosphonate ABC tr
232	20	95.2	241	2	S75347	hypothetical prote
233	20	95.2	241	2	T08536	transfer origin pr
234	20	95.2	241	2	S23004	traL protein - Esc
235	20	95.2	241	2	S59377	probable membrane
236	20	95.2	242	2	A86719	oxidoreductase yhg
237	20	95.2	242	2	A95180	conserved hypothet
238	20	95.2	242	2	E89813	hypothetical prote
239	20	95.2	242	2	D98047	conserved hypothet
240	20	95.2	243	2	AB7478	methyltransferase,
241	20	95.2	244	2	C75605	hypothetical prote
242	20	95.2	244	2	S50685	hypothetical prote
243	20	95.2	245	2	D72258	helicase-related p
244	20	95.2	245	2	F97002	ABC-type multidrug
245	20	95.2	247	2	AD3361	endopeptidase Clp
246	20	95.2	247	2	S33196	osmotin-like prote
247	20	95.2	247	2	S33197	osmotin-like prote
248	20	95.2	248	2	S60940	RSC complex chain

249	20	95.2	248	2	T44822	hypothetical prote
250	20	95.2	249	2	S04674	hypothetical prote
251	20	95.2	249	2	F91274	hypothetical prote
252	20	95.2	249	2	F86115	hypothetical prote
253	20	95.2	249	2	S56415	hypothetical 27.6K
254	20	95.2	250	1	OBNC2	cytochrome-c oxida
255	20	95.2	250	1	C45353	coat protein - app
256	20	95.2	250	2	D71061	hypothetical prote
257	20	95.2	252	1	WMBV2P	29K protein - toba
258	20	95.2	252	2	B87558	hypothetical prote
259	20	95.2	252	2	D81052	acyltransferase, p
260	20	95.2	252	2	AH3267	acetyltransferase
261	20	95.2	252	2	S61142	hypothetical prote
262	20	95.2	254	2	B97582	phosphoribosylamin
263	20	95.2	254	2	AI2802	hypothetical prote
264	20	95.2	254	2	AD3392	phosphoribosylamin
265	20	95.2	254	2	F90408	conserved hypothet
266	20	95.2	255	2	I48235	inhibin beta-B cha
267	20	95.2	255	2	F97235	probable PHP famil
268	20	95.2	256	2	JQ2326	coat protein - Ind
269	20	95.2	256	2	S22589	hypothetical prote
270	20	95.2	256	2	D96743	unknown protein [i
271	20	95.2	257	2	JQ2332	AR1 protein - cass
272	20	95.2	257	2	S39234	gene V2 protein -
273	20	95.2	257	2	T40880	hypothetical prote
274	20	95.2	257	2	AH0835	probable reverse t
275	20	95.2	258	1	QQOMC1	coat protein - cas
276	20	95.2	258	1	VCOMCN	coat protein - cas
277	20	95.2	258	2	S25624	coat protein - cas
278	20	95.2	258	2	S25622	coat protein - cas
279	20	95.2	258	2	S25623	coat protein - cas
280	20	95.2	258	2	T02819	probable membrane
281	20	95.2	258	2	G82868	hypothetical prote
282	20	95.2	259	2	F69678	involved in polyke
283	20	95.2	259	2	T13260	hypothetical prote
284	20	95.2	259	2	D86685	prophage pi1 prote
285	20	95.2	259	2	C86797	prophage pi3 prote
286	20	95.2	259	2	C86757	prophage pi2 prote
287	20	95.2	260	1	A30202	transcription init
288	20	95.2	260	1	QQCVCL	coat protein - tom
289	20	95.2	260	2	A64561	NH(3)-dependent NA
290	20	95.2	260	2	AE0438	hypothetical prote
291	20	95.2	261	2	C86481	30.5K hypothetical
292	20	95.2	262	2	D85753	probable transient
293	20	95.2	262	2	E90865	probable transient
294	20	95.2	264	2	AI1802	an E. coli protein
295	20	95.2	264	2	AB1429	an E. coli protein
296	20	95.2	264	2	F72714	hypothetical prote
297	20	95.2	265	2	E64880	hypothetical prote
298	20	95.2	266	2	G90785	probable acetyltra
299	20	95.2	266	2	E85645	probable acetyltra
300	20	95.2	266	2	G64842	probable hydrolase
301	20	95.2	267	1	A46535	interleukin-2 rece
302	20	95.2	267	2	S76499	hypothetical prote
303	20	95.2	267	2	G90467	conserved hypothet
304	20	95.2	268	2	T04966	hypothetical prote
305	20	95.2	268	2	T38394	probable methyl tr
306	20	95.2	270	2	F64460	hypothetical prote
307	20	95.2	271	2	H83622	hypothetical prote
308	20	95.2	272	2	I48700	gene ox40 protein
309	20	95.2	272	2	B75293	amino acid ABC tra
310	20	95.2	273	2	C72328	transaminase B hom
311	20	95.2	273	2	AD1855	ATP-binding protei
312	20	95.2	273	2	G81952	HemK protein MMA03
313	20	95.2	273	2	AF3017	flagellar motor pr
314	20	95.2	276	2	JC5285	carbonyl reductase
315	20	95.2	276	2	JQ2340	anthranilate synth
316	20	95.2	276	2	F84859	hypothetical prote
317	20	95.2	276	2	AB2037	hypothetical prote
318	20	95.2	278	2	G64393	hypothetical prote
319	20	95.2	279	2	AB0285	probable regulator
320	20	95.2	280	2	AG1856	hypothetical prote
321	20	95.2	281	2	T22670	hypothetical prote
322	20	95.2	282	2	AG3545	dipeptide transpor
323	20	95.2	282	2	T18608	hypothetical prote
324	20	95.2	282	2	E84136	ABC transporter (A
325	20	95.2	282	2	F86396	hypothetical prote
326	20	95.2	283	2	AD2602	conserved hypothet
327	20	95.2	283	2	D97384	endopeptidase-rela
328	20	95.2	284	1	E64938	probable aldehyde
329	20	95.2	284	2	F85788	probable an aldehy
330	20	95.2	284	2	B90940	probable an aldehy
331	20	95.2	284	2	D81075	conserved hypothet
332	20	95.2	284	2	D81866	hypothetical prote
333	20	95.2	284	2	G82319	DnaJ-related prote
334	20	95.2	285	2	E86521	polysaccharide hyd
335	20	95.2	285	2	A72101	polysaccharide hyd
336	20	95.2	285	2	AG1535	transport protein
337	20	95.2	286	2	B83259	hypothetical prote
338	20	95.2	287	1	S49944	actin-capping prot
339	20	95.2	287	2	S26190	nitrogenase (EC 1.
340	20	95.2	287	2	B95016	transcription regu
341	20	95.2	287	2	B86324	protein F14D16.21
342	20	95.2	287	2	D97889	positive transcrip
343	20	95.2	288	2	T45536	moaR protein [impo
344	20	95.2	289	2	S16556	hypothetical prote
345	20	95.2	290	2	AH3495	gamma-D-glutamyl-L
346	20	95.2	291	2	C81943	probable integrase
347	20	95.2	291	2	D89839	hypothetical prote
348	20	95.2	291	2	AC0085	probable membrane
349	20	95.2	292	2	S11870	peroxidase (EC 1.1
350	20	95.2	293	2	AE0372	dihydrodipicolinat
351	20	95.2	293	2	S72970	tyrosine-tRNA liga
352	20	95.2	293	2	AI0404	probable hexulose-
353	20	95.2	294	2	G82483	probable hexulose-
354	20	95.2	294	2	H96719	homeobox gene 13 p
355	20	95.2	295	1	P3WMBB	3a protein - broad
356	20	95.2	297	2	E96736	probable dehydroge
357	20	95.2	297	2	T36724	probable membrane
358	20	95.2	297	2	AE1816	ABC transporter (A
359	20	95.2	298	2	B75612	phosphate ABC tran
360	20	95.2	299	2	D96736	probable dehydroge
361	20	95.2	299	2	C87347	transcription regu
362	20	95.2	299	2	AD3083	transcription regu
363	20	95.2	300	2	F70196	flagellar hook-bas
364	20	95.2	301	2	B64189	glycine cleavage s
365	20	95.2	301	2	E90608	ABC transporter at
366	20	95.2	301	2	T14331	homeotic protein -
367	20	95.2	301	2	B86966	probable oxidoredu
368	20	95.2	302	2	A47126	alsSD operon activ
369	20	95.2	302	2	B49941	devR protein - Myx
370	20	95.2	303	2	D85356	cinnamoyl-CoA redu
371	20	95.2	304	2	G95274	probable gluconola
372	20	95.2	304	2	A95364	protein [imported
373	20	95.2	305	2	B90136	hypothetical prote
374	20	95.2	305	2	C87356	conserved hypothet
375	20	95.2	306	2	C87283	glutaminase A [imp
376	20	95.2	307	2	B81874	hypothetical prote
377	20	95.2	307	2	C81148	YabO/YceC/SfhB fam
378	20	95.2	308	2	C95406	hypothetical prote
379	20	95.2	309	1	EUSMAG	agarase (EC 3.2.1.
380	20	95.2	309	2	F82936	thioredoxin reduct
381	20	95.2	309	2	A83883	sugar transport sy
382	20	95.2	310	2	C72529	hypothetical prote
383	20	95.2	310	2	F96928	ABC-type multidrug
384	20	95.2	310	2	C96929	ABC-type multidrug
385	20	95.2	311	1	LNHU2A	asialoglycoprotein
386	20	95.2	311	2	H95877	hypothetical prote
387	20	95.2	311	2	B72781	hypothetical prote
388	20	95.2	312	2	H86732	hypothetical prote
389	20	95.2	312	2	AB0306	probable membrane
390	20	95.2	312	2	H97342	ABC-type MDR trans
391	20	95.2	312	2	G97260	ABC-type MDR trans
392	20	95.2	312	2	S49612	transposase - Esch
393	20	95.2	313	2	E83616	probable transcrip
394	20	95.2	313	2	G97006	ABC-type multidrug

395	20	95.2	313	2	T12634	homeotic protein -
396	20	95.2	314	2	T32380	hypothetical prote
397	20	95.2	314	2	D97318	ABC-type MDR trans
398	20	95.2	315	2	D96948	2-keto-3-deoxygluc
399	20	95.2	316	2	D75588	transcription regu
400	20	95.2	316	2	B71082	hypothetical prote
401	20	95.2	317	2	F83730	hypothetical prote
402	20	95.2	318	2	H81255	hypothetical prote
403	20	95.2	318	2	S52424	homeodomain protei
404	20	95.2	319	2	AG0114	glutathione syntha
405	20	95.2	319	2	T34728	hypothetical prote
406	20	95.2	320	2	A84651	hypothetical prote
407	20	95.2	321	2	T13539	hypothetical prote
408	20	95.2	323	2	S62736	cathepsin-like cys
409	20	95.2	323	2	JC5691	cysteine proteinas
410	20	95.2	323	2	T22956	hypothetical prote
411	20	95.2	323	2	S59373	cyclin homolog UME
412	20	95.2	324	2	S62735	cathepsin - Choris
413	20	95.2	325	2	E89939	Heat-inducible tra
414	20	95.2	326	2	C91228	hypothetical prote
415	20	95.2	326	2	B86075	hypothetical prote
416	20	95.2	326	2	S40818	hypothetical 36.9K
417	20	95.2	329	2	T18617	hypothetical prote
418	20	95.2	329	2	T30513	hypothetical prote
419	20	95.2	329	2	C96033	probable regulator
420	20	95.2	331	2	D69490	LSU ribosomal prot
421	20	95.2	331	2	D49905	protein secretion
422	20	95.2	331	2	E83457	molybdopterin bios
423	20	95.2	331	2	A81173	conserved hypothet
424	20	95.2	331	2	G81932	probable periplasm
425	20	95.2	331	2	T36969	probable phytoene
426	20	95.2	331	2	H90252	NADH dehydrogenase
427	20	95.2	331	2	A87652	hypothetical prote
428	20	95.2	332	2	AB0808	probable ion-chann
429	20	95.2	334	1	B49888	thioredoxin-disulf
430	20	95.2	334	2	A47180	L-lactate dehydrog
431	20	95.2	335	2	C69279	heat shock protein
432	20	95.2	335	2	G72304	histidinol-phospha
433	20	95.2	336	2	G90336	potassium channel
434	20	95.2	336	2	B90574	ABC transporter pe
435	20	95.2	336	2	D98203	probable transcrip
436	20	95.2	337	2	C81801	ketol-acid reducto
437	20	95.2	337	2	F81066	ketol-acid reducto
438	20	95.2	337	2	AF0086	periplasmic bindin
439	20	95.2	337	2	T33209	hypothetical prote
440	20	95.2	338	2	B90938	asparaginase (EC 3
441	20	95.2	338	2	AC0711	asparaginase (EC 3
442	20	95.2	338	2	XDEC1	asparaginase (EC 3
443	20	95.2	338	2	D72327	heat shock operon
444	20	95.2	338	2	F87267	PDZ domain family
445	20	95.2	338	2	A84593	hypothetical prote
446	20	95.2	338	2	B89625	protein C14H10.1 [
447	20	95.2	339	2	AB3330	ketol-acid reducto
448	20	95.2	339	2	D83308	probable transcrip
449	20	95.2	339	2	H83365	hypothetical prote
450	20	95.2	340	2	T46112	hypothetical prote
451	20	95.2	341	1	KIBE36	thymidine kinase (
452	20	95.2	341	1	KIBE40	thymidine kinase (
453	20	95.2	341	1	KIBE73	thymidine kinase (
454	20	95.2	341	1	KIBEEL	thymidine kinase (
455	20	95.2	341	1	KIBEGK	thymidine kinase (
456	20	95.2	341	2	G89811	hypothetical prote
457	20	95.2	342	2	T22428	hypothetical prote
458	20	95.2	343	2	S09777	hypothetical prote
459	20	95.2	343	2	AF3292	chemotaxis motB pr
460	20	95.2	343	2	B98267	hypothetical prote
461	20	95.2	343	2	T20388	hypothetical prote
462	20	95.2	344	2	S73555	MG415 homolog C12_
463	20	95.2	345	2	B42604	ORF2 complementary
464	20	95.2	345	2	T09018	probable calcium-b
465	20	95.2	345	2	B87375	hypothetical prote
466	20	95.2	345	2	T41473	probable zinc fing
467	20	95.2	345	2	AI2560	hypothetical prote
468	20	95.2	347	1	A44245	alcohol dehydrogen
469	20	95.2	347	1	S51120	alcohol dehydrogen
470	20	95.2	347	2	AG2433	hypothetical prote
471	20	95.2	348	2	A82322	octaprenyl-diphosp
472	20	95.2	349	1	WFPGBB	inhibin beta-B cha
473	20	95.2	350	2	G95382	probable Transport
474	20	95.2	351	2	AE2979	aldo/keto reductas
475	20	95.2	351	2	H98303	hypothetical prote
476	20	95.2	353	1	H71340	membrane lipoprote
477	20	95.2	354	2	T46875	Dbh protein [impor
478	20	95.2	356	2	A91023	isochorismate hydr
479	20	95.2	356	2	G64997	isochorismate synt
480	20	95.2	356	2	B85867	hypothetical prote
481	20	95.2	356	2	T41764	AcMNPV orf18 - Bom
482	20	95.2	356	2	T48354	hypothetical prote
483	20	95.2	357	2	T17471	probable (S)-2-hyd
484	20	95.2	358	2	S41640	ribosomal protein
485	20	95.2	359	2	D83385	hypothetical prote
486	20	95.2	360	1	S30109	D-vitopine dehydro
487	20	95.2	360	2	D69367	DNA topoisomerase
488	20	95.2	360	2	AI2042	hypothetical prote
489	20	95.2	361	2	B84601	cysteine proteinas
490	20	95.2	361	2	E84506	probable N-acetylg
491	20	95.2	361	2	F97068	histidinol-phospha
492	20	95.2	361	2	AG0346	putitive ABC-trans
493	20	95.2	362	2	T19285	hypothetical prote
494	20	95.2	363	2	T40797	60s ribosomal prot
495	20	95.2	363	2	H87277	flagellin modifica
496	20	95.2	364	1	QQBE9	BORF1 protein - hu
497	20	95.2	364	2	T34954	probable UDP-N-ace
498	20	95.2	364	2	T48752	GRR1 related prote
499	20	95.2	365	1	MNXRAH	nonstructural prot
500	20	95.2	366	2	T50468	probable maturase
501	20	95.2	367	1	KIECEG	glutamate 5-kinase
502	20	95.2	367	2	AG0543	glutamate 5-kinase
503	20	95.2	367	2	C85513	gamma-glutamate ki
504	20	95.2	367	2	E90662	gamma-glutamate ki
505	20	95.2	367	2	AE0391	glutamate 5-kinase
506	20	95.2	368	2	JN0718	cysteine proteinas
507	20	95.2	369	2	F83250	histidinol-phospha
508	20	95.2	370	2	I51199	activin beta B sub
509	20	95.2	370	2	A71143	hypothetical prote
510	20	95.2	371	2	C83736	transposase (15) B
511	20	95.2	371	2	AC3210	hypothetical prote
512	20	95.2	373	2	S17955	long-chain-fatty-a
513	20	95.2	373	2	S15161	long-chain-fatty-a
514	20	95.2	374	2	AE2337	hypothetical prote
515	20	95.2	374	2	E64489	hypothetical prote
516	20	95.2	375	2	E70011	potassium channel
517	20	95.2	376	2	T35868	probable dipeptida
518	20	95.2	376	2	B75380	probable undecapre
519	20	95.2	379	2	T10949	cysteine proteinas
520	20	95.2	382	2	T16431	hypothetical prote
521	20	95.2	384	2	T21502	hypothetical prote
522	20	95.2	385	2	F97253	UDP-N-acetylglucos
523	20	95.2	385	2	S43540	YSD83 protein - ye
524	20	95.2	385	2	T36899	probable oxidoredu
525	20	95.2	386	2	T28375	ORF MSV214 SCG gen
526	20	95.2	386	2	T38718	hypothetical prote
527	20	95.2	386	2	AH3624	spermidine/putresc
528	20	95.2	387	2	B83548	hypothetical prote
529	20	95.2	388	2	T26372	hypothetical prote
530	20	95.2	389	2	D90248	DNA topoisomerase
531	20	95.2	389	2	T20604	hypothetical prote
532	20	95.2	390	2	T30395	probable transcrip
533	20	95.2	391	2	T25211	hypothetical prote
534	20	95.2	392	2	A05025	NADH2 dehydrogenas
535	20	95.2	393	2	E95261	serine proteinase
536	20	95.2	394	2	E64045	mannonate dehydrat
537	20	95.2	394	2	AI2546	hypothetical prote
538	20	95.2	394	2	T37520	arginase family pr
539	20	95.2	394	2	T20519	hypothetical prote
540	20	95.2	395	1	Q0ECTR	hypothetical 45.2K

541	20	95.2	395	2	A91129	hypothetical prote
542	20	95.2	395	2	H85973	hypothetical prote
543	20	95.2	395	2	E90896	probable transport
544	20	95.2	395	2	B85721	probable transport
545	20	95.2	395	2	A64908	membrane protein y
546	20	95.2	396	2	AH1426	peptidases homolog
547	20	95.2	396	2	AG1800	peptidases homolog
548	20	95.2	397	2	B98127	serine proteinase
549	20	95.2	397	2	S26731	neuro-D4 protein -
550	20	95.2	398	2	D83601	conserved hypothet
551	20	95.2	398	2	C96680	hypothetical prote
552	20	95.2	399	2	S07630	hypothetical prote
553	20	95.2	399	2	T10314	viral transcriptio
554	20	95.2	399	2	T41786	P47 orf40 - Bombyx
555	20	95.2	400	2	AB1427	carboxypeptidase h
556	20	95.2	400	2	AI1800	carboxypeptidase h
557	20	95.2	400	2	G69370	3-ketoacyl-CoA thi
558	20	95.2	401	2	H72854	viral transcriptio
559	20	95.2	401	2	F95922	probable NDP-hexos
560	20	95.2	403	2	A28451	beta-galactoside a
561	20	95.2	403	2	T39697	DNAJ protein - fis
562	20	95.2	404	2	A99277	hypothetical prote
563	20	95.2	405	2	AF2559	hypothetical prote
564	20	95.2	405	2	AG2534	hypothetical prote
565	20	95.2	406	2	T02291	hypothetical prote
566	20	95.2	407	1	A40150	inhibin beta-B cha
567	20	95.2	407	1	R5FFL1	ribosomal protein
568	20	95.2	407	2	T12085	reverse transcript
569	20	95.2	407	2	T45965	hypothetical prote
570	20	95.2	409	2	F90825	probable integrase
571	20	95.2	410	2	S15163	probable transposa
572	20	95.2	410	2	C44490	retrovirus-related
573	20	95.2	410	2	T47926	hypothetical prote
574	20	95.2	411	2	B41398	inhibin beta-B cha
575	20	95.2	411	2	E64088	ATP-dependent clp
576	20	95.2	412	2	T47321	hypothetical prote
577	20	95.2	414	2	AG2407	site-specific DNA-
578	20	95.2	414	2	S60190	vicilin - Zamia fu
579	20	95.2	416	2	F70017	probable transamin
580	20	95.2	416	2	A85684	probable integrase
581	20	95.2	417	2	AF0251	probable M23/M37 p
582	20	95.2	417	2	T47616	hypothetical prote
583	20	95.2	417	2	B83469	probable chemotaxi
584	20	95.2	418	2	AI1784	UV-damage repair p
585	20	95.2	418	2	AC1409	UV-damage repair p
586	20	95.2	418	2	G83654	transposase (11) B
587	20	95.2	418	2	H83677	transposase (11) B
588	20	95.2	419	2	T47443	hypothetical prote
589	20	95.2	419	2	A70814	probable integral
590	20	95.2	421	2	T51809	succinate-CoA liga
591	20	95.2	423	2	F75635	hypothetical prote
592	20	95.2	423	2	S76384	hypothetical prote
593	20	95.2	423	2	G83573	conserved hypothet
594	20	95.2	423	2	G82145	conserved hypothet
595	20	95.2	423	2	E81010	hemK protein NMB20
596	20	95.2	424	2	JC7558	chromatin assembly
597	20	95.2	424	2	AE0262	conserved hypothet
598	20	95.2	425	2	AH1041	Vi polysaccharide
599	20	95.2	425	2	B36892	Vi polysaccharide
600	20	95.2	425	2	S36112	G1/S transition co
601	20	95.2	425	2	I39181	G1/S transition co
602	20	95.2	425	2	I49367	G1/S transition co
603	20	95.2	426	2	B87078	tyrosyl-tRNA synth
604	20	95.2	426	2	F71299	serine-tRNA ligase
605	20	95.2	427	2	A85789	hypothetical prote
606	20	95.2	427	2	E90940	hypothetical prote
607	20	95.2	427	2	H64938	hypothetical prote
608	20	95.2	427	2	G84375	adenosylhomocystei
609	20	95.2	428	2	B95158	homoserine dehydro
610	20	95.2	428	2	B64081	fucose permease ho
611	20	95.2	428	2	AE0712	conserved hypothet
612	20	95.2	429	2	B71307	conserved hypothet
613	20	95.2	430	2	G86870	histidine-tRNA lig
614	20	95.2	430	2	G72201	conserved hypothet
615	20	95.2	431	2	AE2090	site-specific DNA-
616	20	95.2	431	2	AE0795	isochorismate synt
617	20	95.2	431	2	H81981	probable tyrosine-
618	20	95.2	431	2	AH0860	probable RNA methy
619	20	95.2	432	2	T33118	hypothetical prote
620	20	95.2	433	2	H82139	trigger factor [si
621	20	95.2	433	2	B84923	hypothetical prote
622	20	95.2	433	2	A81793	deoxyribodipyrimid
623	20	95.2	434	2	F82318	tyrosyl-tRNA synth
624	20	95.2	435	2	F98302	edtA monooxygenase
625	20	95.2	435	2	AI2980	nitrilotriacetate
626	20	95.2	435	2	B83958	glucose-inhibited
627	20	95.2	435	2	F89898	glucose-inhibited
628	20	95.2	436	2	C89926	hypothetical prote
629	20	95.2	437	2	T44520	lipopolysaccharide
630	20	95.2	437	2	T44509	Vi polysaccharide
631	20	95.2	438	2	A86502	ATP synthase subun
632	20	95.2	438	2	C72121	ATP synthase, chai
633	20	95.2	438	2	A83823	ATP-dependent RNA
634	20	95.2	438	2	B97712	hypothetical prote
635	20	95.2	439	2	B98024	homoserine dehydro
636	20	95.2	439	2	AF0742	conserved hypothet
637	20	95.2	440	2	D83089	histidinol dehydro
638	20	95.2	440	2	A72405	DNA repair protein
639	20	95.2	440	2	B64090	dicarboxylate tran
640	20	95.2	440	2	D71715	hypothetical prote
641	20	95.2	441	2	G97126	probable Fe-S oxid
642	20	95.2	441	2	G64492	hypothetical prote
643	20	95.2	442	2	T33412	hypothetical prote
644	20	95.2	445	2	H96771	hypothetical prote
645	20	95.2	446	2	S46786	serine-tRNA ligase
646	20	95.2	446	2	B70543	hypothetical prote
647	20	95.2	448	2	G89933	hypothetical prote
648	20	95.2	448	2	E96837	unknown protein T2
649	20	95.2	449	2	A71248	amidophosphoribosy
650	20	95.2	451	2	S76813	hypothetical prote
651	20	95.2	453	2	T38707	probable initiator
652	20	95.2	454	2	T29024	hypothetical prote
653	20	95.2	454	2	T44878	3-oxoacyl-[acyl-ca
654	20	95.2	454	2	T41020	probable Uridine k
655	20	95.2	454	2	F70938	probable fabG4 pro
656	20	95.2	455	2	S16559	cellulase (EC 3.2.
657	20	95.2	455	2	D75043	seryl-tRNA synthet
658	20	95.2	455	2	E71569	probable acyltrans
659	20	95.2	455	2	C97256	uncharacterized pr
660	20	95.2	456	1	E42594	hypothetical prote
661	20	95.2	456	2	T06589	3-methyl-2-oxobuta
662	20	95.2	456	2	T06136	aspartate transami
663	20	95.2	460	2	G71117	serine-tRNA ligase
664	20	95.2	460	2	C83468	conserved hypothet
665	20	95.2	461	2	I49366	G1/S transition co
666	20	95.2	461	2	A57463	p55PIK - mouse
667	20	95.2	462	1	S76100	hypothetical prote
668	20	95.2	462	2	T30164	probable phosphopr
669	20	95.2	462	2	AF0057	probable transport
670	20	95.2	462	2	G01804	interleukin 3-regu
671	20	95.2	463	2	T43344	nuclear receptor N
672	20	95.2	465	2	B81658	2-oxo acid dehydro
673	20	95.2	465	2	E71500	probable lipoamide
674	20	95.2	466	2	AC2696	two component sens
675	20	95.2	466	2	C97478	popq protein [impo
676	20	95.2	467	2	S58233	PopQ protein - Rhi
677	20	95.2	468	2	T20475	hypothetical prote
678	20	95.2	471	2	F72337	hypothetical prote
679	20	95.2	471	2	E86322	hypothetical prote
680	20	95.2	471	2	S61202	probable membrane
681	20	95.2	472	2	C70853	hypothetical prote
682	20	95.2	472	2	AH2925	hypothetical prote
683	20	95.2	472	2	F98356	hypothetical prote
684	20	95.2	473	2	T04225	hypothetical prote
685	20	95.2	473	2	C96516	F16N3.15 [imported
686	20	95.2	474	2	AB2990	polysaccharide bio

687	20	95.2	474	2	G98293	polysaccharide bio
688	20	95.2	476	2	H84291	hypothetical prote
689	20	95.2	478	2	F82175	conserved hypothet
690	20	95.2	479	2	C75513	probable ferredoxi
691	20	95.2	479	2	A25052	fibrinogen beta ch
692	20	95.2	480	1	IBEG	hydroxymethylbilan
693	20	95.2	480	2	A37244	nuclear autoantige
694	20	95.2	480	2	B70446	hypothetical prote
695	20	95.2	481	1	B43674	protein kinase (EC
696	20	95.2	481	1	TVBE17	protein kinase (EC
697	20	95.2	481	2	JE0377	p70 S6 kinase (EC
698	20	95.2	482	2	T22981	hypothetical prote
699	20	95.2	485	2	S52411	ZNF165 protein - h
700	20	95.2	486	2	AI1942	cysteinyl-tRNA syn
701	20	95.2	486	2	AD0373	probable exported
702	20	95.2	487	2	B95275	probable ABC trans
703	20	95.2	488	2	E95382	hypothetical prote
704	20	95.2	489	2	T10283	ecdysteroid UDPglu
705	20	95.2	489	2	S23410	FUN19 protein - ye
706	20	95.2	489	2	B72518	hypothetical prote
707	20	95.2	492	2	B70617	hypothetical prote
708	20	95.2	493	2	T10054	pyruvate kinase (E
709	20	95.2	494	2	S74625	NADH-glutamate syn
710	20	95.2	495	2	S60175	regulatory protein
711	20	95.2	495	2	A48370	nitrogen fixation
712	20	95.2	496	2	C83617	probable aldehyde
713	20	95.2	497	2	G81598	serine hydroxymeth
714	20	95.2	497	2	C81672	serine hydroxymeth
715	20	95.2	497	2	H71516	glycine hydroxymet
716	20	95.2	497	2	G95398	probable ABC trans
717	20	95.2	504	2	JC1306	virion protein hom
718	20	95.2	504	2	T38226	probable serine-th
719	20	95.2	505	2	E87527	exodeoxyribonuclea
720	20	95.2	506	2	F83545	hypothetical prote
721	20	95.2	507	2	H82580	glycerol-3-phospha
722	20	95.2	507	2	D97441	histidyl-trna synt
723	20	95.2	507	2	AF2659	histidyl-tRNA synt
724	20	95.2	507	2	T19067	hypothetical prote
725	20	95.2	508	1	S46008	probable carboxype
726	20	95.2	509	2	T01344	hypothetical prote
727	20	95.2	510	2	T07119	cytochrome P450 CP
728	20	95.2	511	1	E89775	2',3'-cyclic-nucle
729	20	95.2	513	2	C86897	hypothetical prote
730	20	95.2	513	2	S54590	2'-O-ribosyl phosp
731	20	95.2	513	2	A87324	hypothetical prote
732	20	95.2	514	2	S50785	RRN7 protein - yea
733	20	95.2	514	2	D84584	probable protein p
734	20	95.2	514	2	T21286	hypothetical prote
735	20	95.2	515	2	AF3524	hypothetical prote
736	20	95.2	518	1	A27705	alpha-amylase (EC
737	20	95.2	518	2	T43173	myosin heavy chain
738	20	95.2	518	2	D83787	spore germination
739	20	95.2	518	2	AG0784	rtn protein [impor
740	20	95.2	519	2	E86555	serine hydroxymeth
741	20	95.2	519	2	H72067	glycine hydroxymet
742	20	95.2	519	2	T46241	hypothetical prote
743	20	95.2	519	2	T33616	hypothetical prote
744	20	95.2	520	2	E87621	hypothetical prote
745	20	95.2	521	2	C84249	MAPK-activated pro
746	20	95.2	521	2	T37504	hypothetical prote
747	20	95.2	522	2	A33644	signal recognition
748	20	95.2	524	2	F75264	hypothetical prote
749	20	95.2	529	1	W7AD22	early E2A DNA-bind
750	20	95.2	529	1	W7AD25	early E2A DNA-bind
751	20	95.2	529	2	I49504	Tum-P91A antigen -
752	20	95.2	529	2	F82983	conserved hypothet
753	20	95.2	530	2	AC2085	phosphodiesterase/
754	20	95.2	530	2	T48117	hypothetical prote
755	20	95.2	530	2	D70476	DNA helicase - Aqu
756	20	95.2	532	2	B82354	deoxycytidylate de
757	20	95.2	533	2	S33701	homeotic protein D
758	20	95.2	534	2	JC5096	transposase - fung
759	20	95.2	535	2	E96730	hypothetical prote
760	20	95.2	535	2	H83324	probable chemotaxi
761	20	95.2	535	2	B45270	sensory histidine
762	20	95.2	535	2	AE3202	hypothetical prote
763	20	95.2	535	2	A87570	ABC transporter, A
764	20	95.2	536	2	T03034	cytochrome p450 -
765	20	95.2	536	2	T37544	hypothetical serin
766	20	95.2	537	2	B38179	rpoN-2 protein - B
767	20	95.2	537	2	D89889	conserved hypothet
768	20	95.2	537	2	F70597	hypothetical prote
769	20	95.2	539	2	T36660	proteinase - Strep
770	20	95.2	540	2	T47858	hypothetical prote
771	20	95.2	541	2	D96779	probable 3-ketoacy
772	20	95.2	543	1	WZBEM5	gene 19 protein -
773	20	95.2	544	2	T42932	virion tegument pr
774	20	95.2	544	2	AB1178	transport protein
775	20	95.2	545	1	JQ0153	mercury(II) reduct
776	20	95.2	545	2	H83079	hypothetical prote
777	20	95.2	545	2	B82740	two-component syst
778	20	95.2	546	2	G72210	hypothetical prote
779	20	95.2	547	2	B86723	NADH oxidase noxC
780	20	95.2	548	2	AB0365	probable pyridine
781	20	95.2	551	2	T40767	hypothetical prote
782	20	95.2	552	2	S55026	secretion protein
783	20	95.2	553	1	B55483	transcription init
784	20	95.2	553	2	S28713	hypothetical prote
785	20	95.2	553	2	AF2662	two component sens
786	20	95.2	554	2	AI1829	hypothetical prote
787	20	95.2	554	2	E90601	hypothetical prote
788	20	95.2	555	2	D90354	NADH oxidase SSO19
789	20	95.2	555	2	B87670	GGDEF family prote
790	20	95.2	555	2	E97444	probable two-compo
791	20	95.2	555	2	T45351	ferredoxin [import
792	20	95.2	557	2	T24538	hypothetical prote
793	20	95.2	558	2	T32028	hypothetical prote
794	20	95.2	558	2	E83905	hypothetical prote
795	20	95.2	562	2	S46281	P element - fruit
796	20	95.2	564	1	HMIVF7	hemagglutinin prec
797	20	95.2	564	2	S15962	hypothetical prote
798	20	95.2	564	2	T15477	hypothetical prote
799	20	95.2	565	2	T42593	exonuclease (EC 3.
800	20	95.2	567	2	T30799	hypothetical prote
801	20	95.2	567	2	E72156	C6R protein - vari
802	20	95.2	567	2	E35928	hypothetical prote
803	20	95.2	567	2	I36841	E6R protein - vari
804	20	95.2	567	2	T28485	hypothetical prote
805	20	95.2	568	2	AB3341	precorrin-3B C17-m
806	20	95.2	569	2	T01399	hypothetical prote
807	20	95.2	569	2	A43317	germ cell-less pro
808	20	95.2	570	2	C89885	DNA-dependent DNA
809	20	95.2	571	2	H97333	site-specific modi
810	20	95.2	572	2	B84958	proline-tRNA ligas
811	20	95.2	573	2	A82231	transport ATP-bind
812	20	95.2	574	2	B29677	complement C9 prec
813	20	95.2	574	2	T01131	hypothetical prote
814	20	95.2	574	2	B87619	sensor histidine k
815	20	95.2	574	2	S22596	hypothetical prote
816	20	95.2	575	2	T34280	hypothetical prote
817	20	95.2	576	2	A40688	peroxisomal protei
818	20	95.2	578	2	H82872	hypothetical prote
819	20	95.2	579	2	F70000	two-component sens
820	20	95.2	579	2	T16237	hypothetical prote
821	20	95.2	582	2	A43412	semenogelin II pre
822	20	95.2	583	2	T10051	pyruvate kinase (E
823	20	95.2	584	1	S65587	ABC transporter st
824	20	95.2	584	2	S51882	topoisomerase I-re
825	20	95.2	585	2	C82157	hypothetical prote
826	20	95.2	586	2	B70503	probable pyrG prot
827	20	95.2	587	2	E84808	hypothetical prote
828	20	95.2	587	2	T41653	probable transcrip
829	20	95.2	589	2	C87664	methyl-accepting c
830	20	95.2	590	2	S72961	CTP synthase (EC 6
831	20	95.2	590	2	E72015	DNA primase - Chla
832	20	95.2	590	2	D86608	DNA primase [impor

833	20	95.2	591	2	H86267	probable protein p
834	20	95.2	593	2	S51946	pyruvate kinase (E
835	20	95.2	593	2	A64075	DNA primase (EC 2.
836	20	95.2	594	2	A81525	DNA primase, proba
837	20	95.2	599	2	C86161	hypothetical prote
838	20	95.2	600	2	C95113	oligoendopeptidase
839	20	95.2	600	2	B97982	group B oligopepti
840	20	95.2	601	2	AH1003	glutathione-regula
841	20	95.2	601	2	F87548	hypothetical prote
842	20	95.2	605	2	T43974	hypothetical prote
843	20	95.2	608	2	A97685	hypothetical prote
844	20	95.2	608	2	AC2910	hypothetical prote
845	20	95.2	610	2	S19461	probable membrane
846	20	95.2	610	2	T44161	hypothetical prote
847	20	95.2	612	2	S55084	probable membrane
848	20	95.2	614	2	S77221	protein kinase pkn
849	20	95.2	616	2	T38717	probable GTP-bindi
850	20	95.2	617	2	G64972	yegA protein precu
851	20	95.2	617	2	B85833	suppressor of ompF
852	20	95.2	617	2	G90987	suppressor of ompF
853	20	95.2	618	2	T04237	hypothetical prote
854	20	95.2	618	2	AE0770	probable outer mem
855	20	95.2	619	2	C83168	heat shock protein
856	20	95.2	619	2	T44285	heat-shock-cognate
857	20	95.2	619	2	B71559	probable metallopr
858	20	95.2	619	2	H81712	conserved hypothet
859	20	95.2	621	2	S49020	nuclear lamin C pr
860	20	95.2	621	2	F86533	metalloproteinase
861	20	95.2	621	2	A72091	metalloproteinase
862	20	95.2	621	2	D81578	zinc proteinase CP
863	20	95.2	621	2	AB3317	ABC transporter AT
864	20	95.2	623	2	T35377	probable membrane
865	20	95.2	624	1	VCNV87	p87 capsid protein
866	20	95.2	624	2	T10374	p87 capsid protein
867	20	95.2	625	2	C25977	phosphotransferase
868	20	95.2	625	2	S60401	nuclear division p
869	20	95.2	626	1	S77286	excinuclease ABC c
870	20	95.2	628	2	A81999	glucose inhibited
871	20	95.2	631	2	F81227	glucose inhibited
872	20	95.2	631	2	D95348	nitric-oxide reduc
873	20	95.2	632	2	T38617	probable ubiquitin
874	20	95.2	632	2	AF3095	nitric oxide reduc
875	20	95.2	632	2	D98191	trkA-like protein
876	20	95.2	633	2	T14612	hypothetical prote
877	20	95.2	633	2	T28788	hypothetical prote
878	20	95.2	634	2	A83447	heat shock protein
879	20	95.2	634	2	T32324	hypothetical prote
880	20	95.2	637	2	A82738	heat shock protein
881	20	95.2	637	2	D87335	DNA mismatch repai
882	20	95.2	638	2	B83890	hypothetical prote
883	20	95.2	639	2	A60633	tetracycline resis
884	20	95.2	640	2	T26820	hypothetical prote
885	20	95.2	642	2	AI1827	cyclomaltodextrin
886	20	95.2	642	2	D69085	transcription cont
887	20	95.2	642	2	AF2995	conserved hypothet
888	20	95.2	643	2	A43423	dynein 74K chain,
889	20	95.2	644	2	B85758	RNase II, mRNA deg
890	20	95.2	644	2	AB0656	exoribonuclease II
891	20	95.2	644	2	A64877	exoribonuclease II
892	20	95.2	644	2	C90861	RNase II, mRNA deg
893	20	95.2	644	2	AF0272	exoribonuclease II
894	20	95.2	645	2	H86281	protein F10B6.18 [
895	20	95.2	645	2	D90782	probable tail fibe
896	20	95.2	645	2	H85642	probable tail fibe
897	20	95.2	646	2	T34532	hypothetical prote
898	20	95.2	647	2	T49586	related to nif-spe
899	20	95.2	648	2	C69423	DNA helicase homol
900	20	95.2	650	2	G72429	hypothetical prote
901	20	95.2	650	2	D71021	hypothetical prote
902	20	95.2	650	2	D98288	tfua (U39409) [imp
903	20	95.2	650	2	F70974	probable acrA1 pro
904	20	95.2	652	2	B82724	cardiolipin syntha
905	20	95.2	652	2	AD2316	hypothetical prote
906	20	95.2	653	2	B75105	probable DNA helic
907	20	95.2	656	1	QQBET2	UL25 protein - hum
908	20	95.2	656	2	E71080	probable DNA-bindi
909	20	95.2	658	2	E83480	bo-type ubiquinol
910	20	95.2	658	2	D81099	membrane-bound lyt
911	20	95.2	658	2	D81842	probable membrane
912	20	95.2	659	2	AF3489	cytochrome o ubiqu
913	20	95.2	660	2	T41580	probable dna-bindi
914	20	95.2	660	2	T21551	hypothetical prote
915	20	95.2	661	2	S44773	C29E4.5 protein -
916	20	95.2	663	2	H64312	probable DNA helic
917	20	95.2	664	2	B36885	ba-type ubiquinol
918	20	95.2	665	2	AI1947	excinuclease ABC c
919	20	95.2	666	2	A42296	lysozyme 2 (EC 3.2
920	20	95.2	666	2	T05432	hypothetical prote
921	20	95.2	669	2	S74391	excinuclease ABC c
922	20	95.2	669	2	A47302	proline dehydrogen
923	20	95.2	672	2	G82687	bo-type ubiquinol
924	20	95.2	672	2	G69503	signal-transducing
925	20	95.2	676	2	T01084	hypothetical prote
926	20	95.2	676	2	C72749	probable cleavage
927	20	95.2	677	2	T27127	hypothetical prote
928	20	95.2	680	2	T04647	hypothetical prote
929	20	95.2	681	2	S37809	transketolase (EC
930	20	95.2	681	2	A83455	DNA polymerase sub
931	20	95.2	681	2	AF0697	probable type III
932	20	95.2	681	2	T15590	hypothetical prote
933	20	95.2	683	2	S34700	probable purine nu
934	20	95.2	684	2	T47694	probable serine/th
935	20	95.2	686	2	D96611	probable CRK1 prot
936	20	95.2	687	2	E81027	glycyl-tRNA synthe
937	20	95.2	687	2	G81970	probable glycine-t
938	20	95.2	688	2	E86409	hypothetical prote
939	20	95.2	688	2	T32750	hypothetical prote
940	20	95.2	688	2	AB2379	two-component sens
941	20	95.2	689	1	JW0107	very-long-chain ac
942	20	95.2	691	2	T46476	hypothetical prote
943	20	95.2	693	2	T26415	hypothetical prote
944	20	95.2	697	2	D84429	hypothetical prote
945	20	95.2	697	2	F84646	hypothetical prote
946	20	95.2	699	2	T16109	hypothetical prote
947	20	95.2	703	2	T05632	hypothetical prote
948	20	95.2	704	2	B84530	probable RING zinc
949	20	95.2	705	2	D75345	probable 1,4-alpha
950	20	95.2	705	2	T51788	hypothetical prote
951	20	95.2	705	2	A48144	protein kinase CDC
952	20	95.2	706	2	T36176	hypothetical prote
953	20	95.2	706	2	T41024	hypothetical prote
954	20	95.2	708	2	T22377	hypothetical prote
955	20	95.2	709	2	A98299	3-methylcrotonyl-C
956	20	95.2	714	1	S44198	DNA topoisomerase
957	20	95.2	715	2	A34408	peroxidase (EC 1.1
958	20	95.2	717	2	T29816	hypothetical prote
959	20	95.2	719	2	T02154	protein kinase hom
960	20	95.2	719	2	S51739	transcription repr
961	20	95.2	719	2	A81358	hypothetical prote
962	20	95.2	720	2	C84055	cell division prot
963	20	95.2	724	2	D85075	probable athila tr
964	20	95.2	725	2	T21363	hypothetical prote
965	20	95.2	725	2	S64124	probable membrane
966	20	95.2	728	2	AF3299	malate synthase (E
967	20	95.2	729	2	T23474	hypothetical prote
968	20	95.2	730	2	D87365	prolyl oligopeptid
969	20	95.2	733	2	E95335	probable cation tr
970	20	95.2	733	2	D83588	conserved hypothet
971	20	95.2	733	2	F84476	probable Athila re
972	20	95.2	735	2	T40619	neutral trehalase
973	20	95.2	738	2	T38767	probable RNA-bindi
974	20	95.2	741	2	AF2297	hypothetical prote
975	20	95.2	743	2	T26102	hypothetical prote
976	20	95.2	743	2	E87386	periplasmic beta-g
977	20	95.2	744	2	AH2582	malate synthase G
978	20	95.2	744	2	F97364	malate synthase G

979	20	95.2	747	2	F64127	DNA topoisomerase
980	20	95.2	747	2	S01918	sex-specific stora
981	20	95.2	749	2	E95889	probable dehydroge
982	20	95.2	749	2	A45046	basic juvenile hor
983	20	95.2	749	2	C84508	probable cap-bindi
984	20	95.2	751	2	C84367	DNA binding protei
985	20	95.2	751	2	T15230	hypothetical prote
986	20	95.2	752	2	AF0888	topoisomerase IV c
987	20	95.2	752	2	A65089	DNA topoisomerase
988	20	95.2	752	2	A45582	DNA topoisomerase
989	20	95.2	752	2	G85961	DNA topoisomerase
990	20	95.2	752	2	G91116	DNA topoisomerase
991	20	95.2	753	2	S22802	transposase (clone
992	20	95.2	753	2	T01619	hypothetical prote
993	20	95.2	757	2	A48592	transferrin recept
994	20	95.2	760	2	T31556	hypothetical prote
995	20	95.2	760	2	G86373	protein T23E23.14
996	20	95.2	761	2	F82078	topoisomerase IV,
997	20	95.2	761	2	T43788	DNA topoisomerase
998	20	95.2	765	2	T49592	neurofilament trip
999	20	95.2	766	2	A24786	P element transpos
1000	20	95.2	766	2	H83141	probable two-compo

ALIGNMENTS

RESULT 1
AF0749
hypothetical protein STY2156 [imported] - Salmonella enterica subsp. enterica serovar Ty
C;Species: Salmonella enterica subsp. enterica serovar Typhi
A;Note: this species has also been called Salmonella typhi
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AF0749
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
, S.; Moule, S.; O'Gaora, P.
Nature 413, 848-852, 2001
A;Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A;Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A;Reference number: AB0502; MUID:21534947; PMID:11677608
A;Accession: AF0749
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-51 <PAR>
A;Cross-references: UNIPARC:UPI0000059CC5; GB:AL513382; PIDN:CAD05697.1; PID:g16503191;
C;Genetics:
A;Gene: STY2156

Query Match 95.2%; Score 20; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 33 RRLN 36

RESULT 2
S00016
protamine Z1 - smaller spotted catshark
C;Species: Scyliorhinus canicula (smaller spotted catshark, smaller spotted dogfish)
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 09-Jul-2004
C;Accession: S00016; S01463
R;Berlot-Picard, F.; Vodjdani, G.; Doly, J.
Eur. J. Biochem. 165, 553-557, 1987
A;Title: Isolation and characterization of a cDNA clone encoding testis protamine Z1 fro
A;Reference number: S00016; MUID:87246639; PMID:3595601
A;Accession: S00016
A;Molecule type: mRNA
A;Residues: 1-51 <BER>
A;Cross-references: UNIPROT:P08433; UNIPARC:UPI0000171494; GB:X05611; NID:g64302; PIDN:C
R;Gusse, M.; Sautiere, P.; Chauviere, M.; Chevaillier, P.
Biochim. Biophys. Acta 748, 93-98, 1983
A;Title: Extraction, purification and characterization of the sperm protamines of the do
A;Reference number: S01463; MUID:84000513; PMID:6615852
A;Accession: S01463
A;Molecule type: protein
A;Residues: 2-6 <GUS>
A;Cross-references: UNIPARC:UPI000017BF3D
C;Keywords: DNA binding; sperm; testis
F;2-51/Product: protamine Z1 #status predicted <MAT>

Query Match 95.2%; Score 20; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 21 RRLN 24

RESULT 3
D98179
hypothetical protein AGR_L_771 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: D98179
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tu
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: D98179
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-51 <KUR>
A;Cross-references: UNIPROT:Q8U4Z9; UNIPARC:UPI00000D2911; GB:AE007870; PIDN:AAK88958.1
C;Genetics:
A;Gene: AGR_L_771
A;Map position: linear chromosome

Query Match 95.2%; Score 20; DB 2; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 34 RRLN 37

RESULT 4
B84091
hypothetical protein BH3530 [imported] - Bacillus halodurans (strain C-125)
C;Species: Bacillus halodurans
C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C;Accession: B84091
R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fuji, F.; Hir
Nucleic Acids Res. 28, 4317-4331, 2000
A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A;Reference number: A83650; MUID:20512582; PMID:11058132
A;Accession: B84091
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-61 <STO>
A;Cross-references: UNIPROT:Q9K743; UNIPARC:UPI00000C421D; GB:AP001519; GB:BA000004; NI
A;Experimental source: strain C-125
C;Genetics:
A;Gene: BH3530

Query Match 95.2%; Score 20; DB 2; Length 61;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||

```
Db         11 RRLN 14


RESULT 5
B64479
hypothetical protein MJ1435 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: B64479
R;Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake,
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996
A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii
A;Reference number: A64300; MUID:96337999; PMID:8688087
A;Accession: B64479
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-71 <BUL>
A;Cross-references: UNIPROT:Q58830; UNIPARC:UPI000013AACB; GB:U67584; GB:L77117; NID:g15
C;Genetics:
A;Map position: FOR1404404-1404619

  Query Match             95.2%;  Score 20;  DB 2;  Length 71;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         21 RRLN 24


RESULT 6
I39470
apolipoprotein B - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 06-Sep-1996 #sequence_revision 06-Sep-1996 #text_change 21-Jul-2000
C;Accession: I39470
R;Huang, L.S.; Ripps, M.E.; Korman, S.H.; Deckelbaum, R.J.; Breslow, J.L.
J. Biol. Chem. 264, 11394-11400, 1989
A;Title: Hypobetalipoproteinemia due to an apolipoprotein B gene exon 21 deletion derive
A;Reference number: I39469; MUID:89291895; PMID:2567736
A;Accession: I39470
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-72 <RES>
A;Cross-references: UNIPARC:UPI000011F7EE; GB:J04838; NID:g178737; PIDN:AAA53374.1; PID:
C;Genetics:
A;Gene: GDB:APOB
A;Cross-references: GDB:119686; OMIM:107730
A;Map position: 2p24-2p24
A;Introns: 58/3

  Query Match             95.2%;  Score 20;  DB 2;  Length 72;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         61 RRLN 64


RESULT 7
E90729
hypothetical protein ECs0805 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: E90729
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gend
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: E90729
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-73 <HAY>
A;Cross-references: UNIPROT:Q8X3F3; UNIPARC:UPI00000D2997; GB:BA000007; PIDN:BAB34228.1
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs0805

  Query Match             95.2%;  Score 20;  DB 2;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         23 RRLN 26


RESULT 8
B90775
C4-type zinc finger protein (TraR family) [imported] - Escherichia coli (strain O157:H7
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: B90775
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: B90775
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-73 <HAY>
A;Cross-references: UNIPROT:Q9KXG7; UNIPARC:UPI000009C1ED; GB:BA000007; PIDN:BAB34593.1
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs1170

  Query Match             95.2%;  Score 20;  DB 2;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         23 RRLN 26


RESULT 9
F91004
probable C4-type zinc finger protein [imported] - Escherichia coli (strain O157:H7, sub
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: F91004
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: F91004
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-73 <HAY>
A;Cross-references: UNIPROT:Q9EYA1; UNIPARC:UPI00000D2950; GB:BA000007; PIDN:BAB36429.1
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs3006

  Query Match             95.2%;  Score 20;  DB 2;  Length 73;
  Best Local Similarity   100.0%;  Pred. No. 1.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
```

```
              ||||
Db          23 RRLN 26


RESULT 10
E82829
hypothetical protein XF0242 [imported] - Xylella fastidiosa (strain 9a5c)
C;Species: Xylella fastidiosa
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: E82829
R;anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A;Reference number: A82515; MUID:20365717; PMID:10910347
A;Note: for a complete list of authors see reference number A59328 below
A;Accession: E82829
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-74 <SIM>
A;Cross-references: UNIPROT:Q9PGQ6; UNIPARC:UPI00000C2341; GB:AE003878; GB:AE003849; NID
A;Experimental source: strain 9a5c
R;Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
, F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tsuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A;Reference number: A59328
A;Contents: annotation
C;Genetics:
A;Gene: XF0242

  Query Match             95.2%;  Score 20;  DB 2;  Length 74;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         10 RRLN 13


RESULT 11
C64871
hypothetical protein b1240 - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 09-Jul-2004
C;Accession: C64871
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title: The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617; PMID:9278503
A;Accession: C64871
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-76 <BLAT>
A;Cross-references: UNIPROT:P76024; UNIPARC:UPI00000D35B4; GB:AE000222; GB:U00096; NID:g
A;Experimental source: strain K-12, substrain MG1655

  Query Match             95.2%;  Score 20;  DB 2;  Length 76;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         28 RRLN 31


RESULT 12
C91287
hypothetical protein ECs5267 [imported] - Escherichia coli (strain O157:H7, substrain RI
C;Species: Escherichia coli
C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C;Accession: C91287
R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
DNA Res. 8, 11-22, 2001
A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genc
A;Reference number: A99629; MUID:21156231; PMID:11258796
A;Accession: C91287
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-80 <HAY>
A;Cross-references: UNIPROT:Q8X2N0; UNIPARC:UPI00000D2A99; GB:BA000007; PIDN:BAB38690.1,
A;Experimental source: strain O157:H7, substrain RIMD 0509952
C;Genetics:
A;Gene: ECs5267

  Query Match             95.2%;  Score 20;  DB 2;  Length 80;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         31 RRLN 34


RESULT 13
T08595
cysteine proteinase (EC 3.4.22.-) - rape (fragment)
C;Species: Brassica napus (rape)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C;Accession: T08595
R;Buchanan-Wollaston, V.; Ainsworth, C.
Plant Mol. Biol. 33, 821-834, 1997
A;Title: Leaf senescence in Brassica napus: cloning of senescence related genes by subt
A;Reference number: Z16446; MUID:97260386; PMID:9106506
A;Accession: T08595
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-85 <BUC>
A;Cross-references: UNIPROT:Q96352; UNIPARC:UPI00000A98F0; EMBL:U68221; NID:g1527224; P
A;Experimental source: cv. Falcon; senescing leaves
C;Genetics:
A;Gene: LSC790
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase
F;58,59,60/Binding site: substrate (Asn, Ser, Trp) #status predicted

  Query Match             95.2%;  Score 20;  DB 2;  Length 85;
  Best Local Similarity   100.0%;  Pred. No. 2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         27 RRLN 30


RESULT 14
S00180
spermatid protein S1 - smaller spotted catshark
C;Species: Scyliorhinus canicula (smaller spotted catshark, smaller spotted dogfish)
C;Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C;Accession: S00180
R;Chauviere, M.; Martinage, A.; Briand, G.; Sautiere, P.; Chevaillier, P.
Eur. J. Biochem. 169, 105-111, 1987
A;Title: Nuclear basic protein transition during sperm differentiation: amino acid sequ
A;Reference number: S00180; MUID:88055030; PMID:3678229
A;Accession: S00180
A;Molecule type: protein
```

A;Residues: 1-87 <CHA>
A;Cross-references: UNIPROT:P13275; UNIPARC:UPI000013603D
C;Superfamily: sperm histone
C;Keywords: DNA binding; nucleus; phosphoprotein; sperm
F;3,55/Binding site: phosphate (Ser) (covalent) #status experimental

Query Match 95.2%; Score 20; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 70 RRLN 73

RESULT 15
B37475
probable structural component p14 - borna disease virus
C;Species: borna disease virus
C;Date: 16-Feb-1994 #sequence_revision 18-Nov-1994 #text_change 16-Aug-2004
C;Accession: B37475
R;Pyper, J.M.; Richt, J.A.; Brown, L.; Rott, R.; Narayan, O.; Clements, J.E.
Virology 195, 229-238, 1993
A;Title: Genomic organization of the structural proteins of borna disease virus revealed
A;Reference number: A37475; MUID:93303922; PMID:8317098
A;Accession: B37475
A;Status: preliminary
A;Molecule type: nucleic acid
A;Residues: 1-87 <PYP>
A;Cross-references: UNIPROT:Q86622; UNIPARC:UPI00000F867D; GB:S62821; NID:g386390; PIDN:
A;Note: sequence extracted from NCBI backbone (NCBIN:134146, NCBIP:134148)
C;Superfamily: Borna disease virus 24K antigen

Query Match 95.2%; Score 20; DB 2; Length 87;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 14 RRLN 17

RESULT 16
C90392
conserved hypothetical protein [imported] - Sulfolobus solfataricus
C;Species: Sulfolobus solfataricus
C;Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004
C;Accession: C90392
R;She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-
Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P
arrett, R.A.; Ragan, M.A.; Sensen, C.W.; Van der Oost, J.
submitted to GenBank, April 2001
A;Description: Sulfolobus solfataricus complete genome.
A;Reference number: A99139
A;Accession: C90392
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-89 <KUR>
A;Cross-references: UNIPROT:Q97WJ2; UNIPARC:UPI00000646A1; GB:AE006641; NID:g13815524; P
C;Genetics:
A;Gene: SSO10051

Query Match 95.2%; Score 20; DB 2; Length 89;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 37 RRLN 40

RESULT 17
H83003
conserved hypothetical protein PA5148 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 05-Oct-2004
C;Accession: H83003
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; B
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: H83003
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-90 <STO>
A;Cross-references: UNIPARC:UPI00000C5F26; GB:AE004927; GB:AE004091; NID:g9951437; PIDN
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA5148
C;Superfamily: fe(II) trafficking protein YggX

Query Match 95.2%; Score 20; DB 2; Length 90;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 53 RRLN 56

RESULT 18
A95415
hypothetical protein SMa2275 [imported] - Sinorhizobium meliloti (strain 1021) magaplas
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: A95415
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilo
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: A95415
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-91 <KUR>
A;Cross-references: UNIPROT:Q92XL9; UNIPARC:UPI00000CB35D; GB:AE006469; PIDN:AAK65883.1,
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.,
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMa2275
A;Genome: plasmid

Query Match 95.2%; Score 20; DB 2; Length 91;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 60 RRLN 63

RESULT 19
AE0430
hypothetical protein YPO3540 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

```
C;Accession: AE0430
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AE0430
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-92 <KUR>
A;Cross-references: UNIPROT:Q8ZB80; UNIPARC:UPI00000CDA3A; GB:AL590842; PIDN:CAC92769.1;
C;Genetics:
A;Gene: YPO3540

  Query Match             95.2%;  Score 20;  DB 2;  Length 92;
  Best Local Similarity   100.0%;  Pred. No. 2.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         87 RRLN 90


RESULT 20
RBBP22
abc1 protein - phage P22
C;Species: phage P22
A;Note: host Salmonella typhimurium
C;Date: 31-Mar-1989 #sequence_revision 31-Mar-1989 #text_change 09-Jul-2004
C;Accession: A27841
R;Murphy, K.C.; Fenton, A.C.; Poteete, A.R.
Virology 160, 456-464, 1987
A;Title: Sequence of the bacteriophage P22 anti-recBCD (abc) genes and properties of P22
A;Reference number: A94366; MUID:88019195; PMID:3660589
A;Accession: A27841
A;Molecule type: DNA
A;Residues: 1-94 <MUR>
A;Cross-references: UNIPROT:P11190; UNIPARC:UPI0000003A83; GB:J02471; GB:M17737; NID:g21
A;Note: the authors translated the codon TTA for residues 4 and 91 as Lys, ACG for resid
C;Genetics:
A;Gene: abc1
C;Superfamily: phage P22 abc1 protein

  Query Match             95.2%;  Score 20;  DB 1;  Length 94;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 21
AD0229
hypothetical protein YPO1879 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AD0229
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AD0229
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-95 <KUR>
A;Cross-references: UNIPROT:Q8ZF39; UNIPARC:UPI00000DC7AF; GB:AL590842; PIDN:CAC90696.1;
C;Genetics:
A;Gene: YPO1879

  Query Match             95.2%;  Score 20;  DB 2;  Length 95;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         84 RRLN 87


RESULT 22
B26074
cysteine proteinase (EC 3.4.22.-) 13 - papaya (fragment)
C;Species: Carica papaya (papaya)
C;Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
C;Accession: B26074
R;McKee, R.A.; Adams, S.; Matthews, J.A.; Smith, C.J.; Smith, H.
Biochem. J. 237, 105-110, 1986
A;Title: Molecular cloning of two cysteine proteinases from paw-paw (Carica papaya).
A;Reference number: A90332; MUID:87099799; PMID:3541893
A;Accession: B26074
A;Molecule type: mRNA
A;Residues: 1-96 <MCK>
A;Cross-references: UNIPROT:P05993; UNIPARC:UPI0000131300; GB:X03971; GB:M25360; NID:g1
C;Superfamily: papain
C;Keywords: cysteine proteinase; hydrolase

  Query Match             95.2%;  Score 20;  DB 2;  Length 96;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         27 RRLN 30


RESULT 23
S51929
homeotic protein CHB5 - carrot
C;Species: Daucus carota (carrot)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 05-Oct-2004
C;Accession: S51929
R;Kawahara, R.; Komamine, A.; Fukuda, H.
Plant Mol. Biol. 27, 155-164, 1995
A;Title: Isolation and characterization of homeobox-containing genes of carrot.
A;Reference number: S51925; MUID:95169997; PMID:7865785
A;Accession: S51929
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-96 <KAW>
A;Cross-references: UNIPARC:UPI000017A2CD
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;2-58/Domain: homeobox homology <HOX>

  Query Match             95.2%;  Score 20;  DB 2;  Length 96;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          6 RRLN 9


RESULT 24
S51928
homeotic protein CHB4 - carrot
C;Species: Daucus carota (carrot)
C;Date: 14-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 05-Oct-2004
C;Accession: S51928
R;Kawahara, R.; Komamine, A.; Fukuda, H.
Plant Mol. Biol. 27, 155-164, 1995
A;Title: Isolation and characterization of homeobox-containing genes of carrot.
A;Reference number: S51925; MUID:95169997; PMID:7865785
A;Accession: S51928
```

```
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-96 <KAW>
A;Cross-references: UNIPROT:Q43428; UNIPARC:UPI000017A2CC
C;Keywords: DNA binding; homeobox; nucleus; transcription regulation
F;2-58/Domain: homeobox homology <HOX>

  Query Match             95.2%;  Score 20;  DB 2;  Length 96;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          6 RRLN 9


RESULT 25
S62346
L71-5 protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: S62346
R;Wright, L.G.; Chen, T.; Thummel, C.S.; Guild, G.M.
J. Mol. Biol. 255, 387-400, 1996
A;Title: Molecular characterization of the 71E late puff in Drosophila melanogaster reve
A;Reference number: S62333; MUID:96152797; PMID:8568884
A;Accession: S62346
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-98 <WRI>
A;Cross-references: UNIPROT:Q24074; UNIPARC:UPI0000083ECE; EMBL:U24244; NID:g775234; PID
C;Genetics:
A;Gene: L71-5
A;Cross-references: FlyBase:FBgn0004593
C;Superfamily: L71-10 protein

  Query Match             95.2%;  Score 20;  DB 2;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         37 RRLN 40


RESULT 26
A39437
exopolysaccharide synthesis protein exoX - Rhizobium meliloti
C;Species: Rhizobium meliloti
C;Date: 21-Feb-1992 #sequence_revision 21-Feb-1992 #text_change 09-Jul-2004
C;Accession: A39437
R;Reed, J.W.; Capage, M.; Walker, G.C.
J. Bacteriol. 173, 3776-3788, 1991
A;Title: Rhizobium meliloti exoG and exoJ mutations affect the exoX-exoY system for modu
A;Reference number: A39437; MUID:91267943; PMID:2050634
A;Accession: A39437
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-98 <REE>
A;Cross-references: UNIPROT:Q02730; UNIPARC:UPI000012A387; GB:M61751; NID:g152185; PIDN:

  Query Match             95.2%;  Score 20;  DB 2;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         88 RRLN 91


RESULT 27
F95975
posttranscription regulator, repressor protein [imported] - Sinorhizobium meliloti (stra
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: F95975
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing end
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: F95975
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-98 <KUR>
A;Cross-references: UNIPROT:Q02730; UNIPARC:UPI000012A387; GB:AL591985; PIDN:CAC49470.1
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: exoX; SMb20947
A;Genome: plasmid

  Query Match             95.2%;  Score 20;  DB 2;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         88 RRLN 91


RESULT 28
G96029
hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid pSym
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: G96029
R;Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Herna
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001
A;Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing end
A;Reference number: A95842; MUID:21396508; PMID:11481431
A;Accession: G96029
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-98 <KUR>
A;Cross-references: UNIPROT:Q92TK8; UNIPARC:UPI00000CB8B2; GB:AL591985; PIDN:CAC49903.1
A;Experimental source: strain 1021, megaplasmid pSymB
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMb20591
A;Genome: plasmid

  Query Match             95.2%;  Score 20;  DB 2;  Length 98;
  Best Local Similarity   100.0%;  Pred. No. 2.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         33 RRLN 36
```

```
RESULT 29
S62333
L71-1 protein - fruit fly (Drosophila melanogaster)
C;Species: Drosophila melanogaster
C;Date: 19-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: S62333; S62342
R;Wright, L.G.; Chen, T.; Thummel, C.S.; Guild, G.M.
J. Mol. Biol. 255, 387-400, 1996
A;Title: Molecular characterization of the 71E late puff in Drosophila melanogaster reve
A;Reference number: S62333; MUID:96152797; PMID:8568884
A;Accession: S62333
A;Molecule type: DNA
A;Residues: 1-100 <WRI>
A;Cross-references: UNIPROT:Q27316; UNIPARC:UPI000008405D; EMBL:U23836; NID:g939996; PID
A;Accession: S62342
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-100 <WRW>
A;Cross-references: UNIPARC:UPI000008405D; EMBL:U24095; NID:g775224; PIDN:AAA65109.1; PI
C;Genetics:
A;Gene: L71-1
A;Cross-references: FlyBase:FBgn0004588
A;Introns: 12/1; 79/1
C;Superfamily: L71-10 protein

  Query Match             95.2%;  Score 20;  DB 2;  Length 100;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         37 RRLN 40


RESULT 30
T50981
hypothetical protein B24P7.360 [imported] - Neurospora crassa (fragment)
C;Species: Neurospora crassa
C;Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 02-Sep-2000
C;Accession: T50981
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, July 2000
A;Reference number: Z25286
A;Accession: T50981
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-102 <SCH>
A;Cross-references: UNIPARC:UPI000017B492; EMBL:AL389890; GSPDB:GN00116; NCSP:B24P7.360
A;Experimental source: BAC clone B24P7; strain OR74A
C;Genetics:
A;Gene: NCSP:B24P7.360
A;Map position: 6

  Query Match             95.2%;  Score 20;  DB 2;  Length 102;
  Best Local Similarity   100.0%;  Pred. No. 2.4e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         55 RRLN 58


RESULT 31
E82425
hypothetical protein VCA0721 [imported] - Vibrio cholerae (strain N16961 serogroup O1)
C;Species: Vibrio cholerae
C;Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C;Accession: E82425
R;Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A;Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A;Reference number: A82035; MUID:20406833; PMID:10952301
A;Accession: E82425
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-103 <HEI>
A;Cross-references: UNIPROT:Q9KLM2; UNIPARC:UPI00000C3612; GB:AE004401; GB:AE003853; NID
A;Experimental source: serogroup O1; strain N16961; biotype El Tor
C;Genetics:
A;Gene: VCA0721
A;Map position: 2

  Query Match             95.2%;  Score 20;  DB 2;  Length 103;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         36 RRLN 39


RESULT 32
S43813
hypothetical protein 39.1 - phage SPP1
C;Species: phage SPP1
C;Date: 27-Jan-1995 #sequence_revision 27-Jan-1995 #text_change 17-Mar-1999
C;Accession: S43813; S41187
R;Pedre, X.; Weise, F.; Chai, S.; Lueder, G.; Alonso, J.C.
J. Mol. Biol. 236, 1324-1340, 1994
A;Title: Analysis of cis and trans acting elements required for the initiation of DNA r
A;Reference number: S43798; MUID:94172631; PMID:8126723
A;Accession: S43813
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-103 <PED>
A;Cross-references: UNIPARC:UPI000017A850; EMBL:X67865

  Query Match             95.2%;  Score 20;  DB 2;  Length 103;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         95 RRLN 98


RESULT 33
T06820
lipid transfer protein - garden pea
C;Species: Pisum sativum (garden pea)
C;Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06820
R;Reddy, M.K.; Nair, S.; Tewari, K.K.
submitted to the EMBL Data Library, August 1997
A;Reference number: Z15832
A;Accession: T06820
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-105 <RED>
A;Cross-references: UNIPROT:O24309; UNIPARC:UPI000009D5F2; EMBL:Y14560; PIDN:CAA74892.1
C;Genetics:
A;Gene: NLTP
C;Superfamily: phospholipid transfer protein

  Query Match             95.2%;  Score 20;  DB 2;  Length 105;
  Best Local Similarity   100.0%;  Pred. No. 2.5e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         44 RRLN 47


RESULT 34
```

F84479
En/Spm-like transposon protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: F84479
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: F84479
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-108 <STO>
A;Cross-references: UNIPROT:Q9ZU61; UNIPARC:UPI00000A1DD3; GB:AE002093; NID:g4262212; PI
C;Genetics:
A;Gene: At2g06730
A;Map position: 2

Query Match 95.2%; Score 20; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 97 RRLN 100

RESULT 35
CABO
cytochrome-c oxidase (EC 1.9.3.1) chain Va [validated] - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 30-Nov-1979 #sequence_revision 08-Oct-1981 #text_change 09-Jul-2004
C;Accession: A00493; D29968
R;Tanaka, M.; Haniu, M.; Yasunobu, K.T.; Yu, C.A.; Yu, L.; Wei, Y.H.; King, T.E.
J. Biol. Chem. 254, 3879-3885, 1979
A;Title: Amino acid sequence of subunit V of bovine heart cytochrome oxidase, the heme a
A;Reference number: A00493; MUID:79173095; PMID:220224
A;Accession: A00493
A;Molecule type: protein
A;Residues: 1-109 <TAN>
A;Cross-references: UNIPROT:P00426; UNIPARC:UPI0000112899
A;Experimental source: heart
R;Yanamura, W.; Zhang, Y.Z.; Takamiya, S.; Capaldi, R.A.
Biochemistry 27, 4909-4914, 1988
A;Title: Tissue-specific differences between heart and liver cytochrome c oxidase.
A;Reference number: A90531; MUID:89000697; PMID:2844245
A;Accession: D29968
A;Molecule type: protein
A;Residues: 1-34 <YAN>
A;Cross-references: UNIPARC:UPI000017217B
A;Experimental source: liver
R;Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Shinzawa-itoh,
submitted to the Brookhaven Protein Data Bank, April 1996
A;Reference number: A67451; PDB:1OCC
A;Contents: annotation; X-ray crystallography, 2.8 angstroms, residues 1-109
R;Tsukihara, T.; Aoyama, H.; Yamashita, E.; Tomizaki, T.; Yamaguchi, H.; Sinzawa-Itoh, K
Science 272, 1136-1144, 1996
A;Title: The whole structure of the 13-subunit oxidized cytochrome c oxidase at 2.8 angs
A;Reference number: A57981; MUID:96216288; PMID:8638158
A;Contents: annotation; X-ray crystallography, 2.8 angstroms
C;Genetics:
A;Genome: nuclear
C;Complex: part of a 13 chain complex spanning the inner mitochondrial membrane and cons
, Vb (see PIR:OGBO6A), VIa (see PIR:OGBO6), VIb (see PIR:OGBO7), VIc (see PIR:OGBO6C), V
dimers within the mitochondrial inner-membrane
C;Function:
A;Description: the cytochrome-c oxidase complex catalyzes the oxidation of four molecule
ns from the mitochondrial matrix producing two molecules of water and lowering the conce
A;Pathway: oxidative phosphorylation; respiratory chain
A;Note: the role of chain Va is not clear
C;Superfamily: mammalian cytochrome-c oxidase chain Va
C;Keywords: electron transfer; membrane-associated complex; mitochondrial inner membran
F;1-109/Product: cytochrome-c oxidase chain Va #status experimental <MAT>
F;1-109/Domain: mitochondrial matrix #status experimental <MM1>

Query Match 95.2%; Score 20; DB 1; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 56 RRLN 59

RESULT 36
AF0301
conserved hypothetical protein YPO2469 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AF0301
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AF0301
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-110 <KUR>
A;Cross-references: UNIPROT:Q8ZDT7; UNIPARC:UPI00000CD916; GB:AL590842; PIDN:CAC91274.1
C;Genetics:
A;Gene: YPO2469

Query Match 95.2%; Score 20; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 4 RRLN 7

RESULT 37
F71149
hypothetical protein PH0408 - Pyrococcus horikoshii
C;Species: Pyrococcus horikoshii
C;Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004
C;Accession: F71149
R;Kawarabayasi, Y.; Sawada, M.; Horikawa, H.; Haikawa, Y.; Hino, Y.; Yamamoto, S.; Seki
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguch
DNA Res. 5, 55-76, 1998
A;Title: Complete sequence and gene organization of the genome of a hyper-thermophilic
A;Reference number: A71000; MUID:98344137; PMID:9679194
A;Accession: F71149
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-112 <KAW>
A;Cross-references: UNIPROT:O58145; UNIPARC:UPI0000062DE3; GB:AP000002; NID:g3236129; P
A;Experimental source: strain OT3
A;Note: this accession replaces an interim accession for a sequence replaced by GenBank
C;Genetics:
A;Gene: PH0408

Query Match 95.2%; Score 20; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 RRLN 4
||||
Db 21 RRLN 24

RESULT 38

```
E83073
hypothetical protein PA4575 [imported] - Pseudomonas aeruginosa (strain PAO1)
C;Species: Pseudomonas aeruginosa
C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C;Accession: E83073
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho
A;Reference number: A82950; MUID:20437337; PMID:10984043
A;Accession: E83073
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <STO>
A;Cross-references: UNIPROT:Q9HVK9; UNIPARC:UPI00000C5D6D; GB:AE004871; GB:AE004091; NID
A;Experimental source: strain PAO1
C;Genetics:
A;Gene: PA4575

  Query Match             95.2%;  Score 20;  DB 2;  Length 113;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         94 RRLN 97


RESULT 39
AD2868
hypothetical protein Atu2374 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AD2868
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AD2868
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <KUR>
A;Cross-references: UNIPROT:Q8UCV9; UNIPARC:UPI00001646FB; GB:AE008688; PIDN:AAL43362.1;
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu2374
A;Map position: circular chromosome

  Query Match             95.2%;  Score 20;  DB 2;  Length 113;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         99 RRLN 102


RESULT 40
AI0294
probable membrane protein YPO2417 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AI0294
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AI0294
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-114 <KUR>
A;Cross-references: UNIPROT:Q93PD5; UNIPARC:UPI00000D0133; GB:AL590842; PIDN:CAC91221.1;
C;Genetics:
A;Gene: YPO2417

  Query Match             95.2%;  Score 20;  DB 2;  Length 114;
  Best Local Similarity   100.0%;  Pred. No. 2.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         34 RRLN 37


RESULT 41
AF0052
hypothetical protein YPO0423 [imported] - Yersinia pestis (strain CO92)
C;Species: Yersinia pestis
C;Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C;Accession: AF0052
R;Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A;Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A;Reference number: AB0001; MUID:21470413; PMID:11586360
A;Accession: AF0052
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-115 <KUR>
A;Cross-references: UNIPROT:Q8ZIR6; UNIPARC:UPI00000DC920; GB:AL590842; PIDN:CAC89281.1
C;Genetics:
A;Gene: YPO0423

  Query Match             95.2%;  Score 20;  DB 2;  Length 115;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         30 RRLN 33


RESULT 42
S64490
hypothetical protein YGR176w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein G7131
C;Species: Saccharomyces cerevisiae
C;Date: 17-May-1996 #sequence_revision 17-May-1996 #text_change 09-Jul-2004
C;Accession: S64490; PS0248
R;Hebling, U.; Hofmann, B.; Delius, H.
submitted to the Protein Sequence Database, May 1996
A;Reference number: S64003
A;Accession: S64490
A;Molecule type: DNA
A;Residues: 1-115 <HEB>
A;Cross-references: UNIPROT:P32475; UNIPARC:UPI000013AF8D; EMBL:Z72961; NID:g1323311; P
A;Experimental source: strain S288C
R;Jandrositz, A.; Turnowsky, F.; Hoegenauer, G.
Gene 107, 155-160, 1991
A;Title: The gene encoding squalene epoxidase from Saccharomyces cerevisiae: cloning an
A;Reference number: JH0490; MUID:92077427; PMID:1743514
A;Accession: PS0248
A;Molecule type: DNA
A;Residues: 1-55,'NSPR' <JAN>
A;Cross-references: UNIPARC:UPI000017CCAC; GB:M64994
C;Genetics:
A;Cross-references: SGD:S0003408
A;Map position: 7R
```

```
  Query Match             95.2%;  Score 20;  DB 2;  Length 115;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         95 RRLN 98


RESULT 43
T18073
hypothetical protein A571R - Chlorella virus PBCV-1
C;Species: Chlorella virus PBCV-1
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T18073
R;Graves, M.V.; Van Etten, J.L.
submitted to the EMBL Data Library, May 1999
A;Reference number: Z18806
A;Accession: T18073
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-116 <GRA>
A;Cross-references: UNIPROT:O41053; UNIPARC:UPI00000F3C73; EMBL:U42580; NID:g4028896; PI
A;Experimental source: specific host Chlorella strain NC64A
C;Genetics:
A;Note: A571R

  Query Match             95.2%;  Score 20;  DB 2;  Length 116;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         96 RRLN 99


RESULT 44
H97644
hypothetical protein AGR_C_4309 [imported] - Agrobacterium tumefaciens (strain C58, Cere
C;Species: Agrobacterium tumefaciens
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: H97644
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: H97644
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-118 <KUR>
A;Cross-references: UNIPROT:Q8UCV9; UNIPARC:UPI00000D1E9B; GB:AE007869; PIDN:AAK88113.1;
C;Genetics:
A;Gene: AGR_C_4309
A;Map position: circular chromosome

  Query Match             95.2%;  Score 20;  DB 2;  Length 118;
  Best Local Similarity   100.0%;  Pred. No. 2.8e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db        104 RRLN 107


RESULT 45
B72663
hypothetical protein APE0730 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: B72663
R;Kawarabayasi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.;
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropy
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: B72663
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-121 <KAW>
A;Cross-references: UNIPROT:Q9YE40; UNIPARC:UPI000005DC77; DDBJ:AP000060; NID:g5104188;
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE0730
C;Superfamily: Aeropyrum pernix hypothetical protein APE0730

  Query Match             95.2%;  Score 20;  DB 2;  Length 121;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         67 RRLN 70


RESULT 46
G70074
hypothetical protein yxeE - Bacillus subtilis
C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: G70074
R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte
C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F
Koetter, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee
y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell
 Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,
A;Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; MUID:98044033; PMID:9384377
A;Accession: G70074
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-121 <KUN>
A;Cross-references: UNIPROT:P54944; UNIPARC:UPI0000060CB2; GB:Z99124; GB:AL009126; NID:
A;Experimental source: strain 168
C;Genetics:
A;Gene: yxeE

  Query Match             95.2%;  Score 20;  DB 2;  Length 121;
  Best Local Similarity   100.0%;  Pred. No. 2.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         94 RRLN 97


RESULT 47
G01477
ribosomal protein L35 - human
C;Species: Homo sapiens (man)
C;Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 09-Jul-2004
C;Accession: G01477
R;Patel, S.K.
submitted to the EMBL Data Library, July 1994
A;Reference number: G07290
A;Accession: G01477
```

```
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-123 <PAT>
A;Cross-references: UNIPROT:P42766; UNIPARC:UPI000015A4DD; EMBL:U12465; NID:g562073; PID
C;Superfamily: rat ribosomal protein L35

  Query Match             95.2%;  Score 20;  DB 2;  Length 123;
  Best Local Similarity   100.0%;  Pred. No. 3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         93 RRLN 96


RESULT 48
B95328
hypothetical protein SMa0983 [imported] - Sinorhizobium meliloti (strain 1021) magaplasm
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: B95328
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bows
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium melilot
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: B95328
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-123 <KUR>
A;Cross-references: UNIPROT:Q92ZF6; UNIPARC:UPI00000CB109; GB:AE006469; PIDN:AAK65188.1;
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: SMa0983
A;Genome: plasmid

  Query Match             95.2%;  Score 20;  DB 2;  Length 123;
  Best Local Similarity   100.0%;  Pred. No. 3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         55 RRLN 58


RESULT 49
T37118
probable transposase, truncated [imported] - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 15-Sep-2000
C;Accession: T37118
R;Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, August 1999
A;Reference number: Z21588
A;Accession: T37118
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-124 <SAU>
A;Cross-references: UNIPARC:UPI00001795FB; EMBL:AL109950; PIDN:CAB52969.1; GSPDB:GN0007
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCJ4.35
C;Superfamily: Streptomyces coelicolor probable transposase SC6G9.36c

  Query Match             95.2%;  Score 20;  DB 2;  Length 124;
  Best Local Similarity   100.0%;  Pred. No. 3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db        116 RRLN 119


RESULT 50
AF2757
conserved hypothetical protein Atu1470 [imported] - Agrobacterium tumefaciens (strain C5
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AF2757
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
erage, G.; Gillet, W.; Grant, C.; Guenthner, D.; Kutyavin, T.; Levy, R.; Li, M.; McClel
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AF2757
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-125 <KUR>
A;Cross-references: UNIPROT:Q8UFC8; UNIPARC:UPI00001283E7; GB:AE008688; PIDN:AAL42476.1,
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu1470
A;Map position: circular chromosome
C;Superfamily: hypothetical protein MJ1523

  Query Match             95.2%;  Score 20;  DB 2;  Length 125;
  Best Local Similarity   100.0%;  Pred. No. 3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         57 RRLN 60


Search completed: January 25, 2006, 18:41:53
Job time : 38.5 secs
```

OM protein - protein search, using sw model

Run on: January 25, 2006, 18:33:23 ; Search time 63 Seconds
(without alignments)
55.994 Million cell updates/sec

Title: US-10-771-242-295
Perfect score: 21
Sequence: 1 RRLNX 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 1000 summaries

Database : UniProt_05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	95.2	16	2	Q23912_DICDI	Q23912 dictyosteli
2	20	95.2	27	2	Q5BX77_SCHJA	Q5bx77 schistosoma
3	20	95.2	33	2	Q50L67_DROBP	Q50167 drosophila
4	20	95.2	33	2	Q50L79_9DIPT	Q50179 drosophila
5	20	95.2	33	2	Q50L83_9DIPT	Q50183 drosophila
6	20	95.2	33	2	Q50L86_9DIPT	Q50186 drosophila
7	20	95.2	33	2	Q50L89_9DIPT	Q50189 drosophila
8	20	95.2	34	2	Q50L77_DROAN	Q50177 drosophila
9	20	95.2	35	2	Q4ZG64_HUMAN	Q4zg64 homo sapien
10	20	95.2	35	2	Q888U8_PSESM	Q888u8 pseudomonas
11	20	95.2	35	2	Q4SVA1_TETNG	Q4sva1 tetraodon n
12	20	95.2	36	2	Q4Z5P5_PLABE	Q4z5p5 plasmodium
13	20	95.2	36	2	Q7ULM0_RHOBA	Q7ulm0 rhodopirell
14	20	95.2	37	2	Q7SHZ1_NEUCR	Q7shz1 neurospora
15	20	95.2	37	2	Q8F419_LEPIN	Q8f419 leptospira
16	20	95.2	39	2	Q4TTJ0_STRPU	Q4ttj0 strongyloce
17	20	95.2	39	2	Q4TTJ2_STRPU	Q4ttj2 strongyloce
18	20	95.2	39	2	Q4TTJ4_STRPU	Q4ttj4 strongyloce
19	20	95.2	39	2	Q4TTJ8_STRPU	Q4ttj8 strongyloce
20	20	95.2	39	2	Q4TTK1_STRPU	Q4ttk1 strongyloce
21	20	95.2	39	2	Q66546_9GAMA	Q66546 human herpe
22	20	95.2	40	2	Q4TTG6_STRDR	Q4ttg6 strongyloce
23	20	95.2	40	2	Q663L4_YERPS	Q663l4 yersinia ps
24	20	95.2	43	2	Q7RL48_PLAYO	Q7rl48 plasmodium
25	20	95.2	43	2	Q8KEJ5_CHLTE	Q8kej5 chlorobium
26	20	95.2	44	2	Q5C5H2_SCHJA	Q5c5h2 schistosoma
27	20	95.2	44	2	Q72AJ8_DESVH	Q72aj8 desulfovibr
28	20	95.2	47	2	Q99IR3_9GEMI	Q99ir3 tomato leaf
29	20	95.2	47	2	Q99IR5_9GEMI	Q99ir5 tomato leaf
30	20	95.2	48	2	Q99IR1_9GEMI	Q99ir1 soybean cri
31	20	95.2	50	1	PRTZ1_SCYCA	P08433 scyliorhinu
32	20	95.2	50	2	Q6DTN0_CANGA	Q6dtn0 candida gla
33	20	95.2	50	2	Q831H1_ENTFA	Q831h1 enterococcu
34	20	95.2	50	2	Q4T9A7_TETNG	Q4t9a7 tetraodon n
35	20	95.2	51	2	Q9ZSF6_9SOLA	Q9zsf6 lycium ande
36	20	95.2	51	2	Q8RTK9_LEUME	Q8rtk9 leuconostoc
37	20	95.2	51	2	Q8U4Z9_AGRT5	Q8u4z9 agrobacteri
38	20	95.2	51	2	Q57N53_SALCH	Q57n53 salmonella
39	20	95.2	51	2	Q8VKJ1_MYCTU	Q8vkj1 mycobacteri
40	20	95.2	51	2	Q5PI12_SALPA	Q5pi12 salmonella
41	20	95.2	51	2	Q8Z5T2_SALTI	Q8z5t2 salmonella
42	20	95.2	51	2	Q8ZNT9_SALTY	Q8znt9 salmonella
43	20	95.2	52	2	Q96I42_HUMAN	Q96i42 homo sapien
44	20	95.2	52	2	Q54BT9_DICDI	Q54bt9 dictyosteli
45	20	95.2	53	2	Q9ZSG7_9SOLA	Q9zsg7 lycium ande
46	20	95.2	54	2	Q8EDQ1_SHEON	Q8edq1 shewanella
47	20	95.2	54	2	Q8DI57_SYNEL	Q8di57 synechococc
48	20	95.2	54	2	Q9E3W1_TYLCV	Q9e3w1 tomato yell
49	20	95.2	54	2	Q4GXV3_9GEMI	Q4gxv3 tomato gemi
50	20	95.2	55	2	Q4GXV7_9GEMI	Q4gxv7 tomato gemi
51	20	95.2	56	2	Q5JVT5_HUMAN	Q5jvt5 homo sapien
52	20	95.2	56	2	Q4GXV1_9GEMI	Q4gxv1 tomato gemi
53	20	95.2	58	2	Q6UU04_ORYSA	Q6uu04 oryza sativ
54	20	95.2	58	2	Q6FD70_ACIAD	Q6fd70 acinetobact
55	20	95.2	59	2	Q6JCV6_XANCG	Q6jcv6 xanthomonas
56	20	95.2	59	2	Q4GXV9_9GEMI	Q4gxv9 tomato gemi
57	20	95.2	60	2	Q64A06_9ARCH	Q64a06 uncultured
58	20	95.2	60	2	Q4VWT6_9UROC	Q4vwt6 oikopleura
59	20	95.2	60	2	Q4LVY7_9BURK	Q4lvy7 burkholderi
60	20	95.2	60	2	Q8XV30_RALSO	Q8xv30 ralstonia s
61	20	95.2	60	2	Q62GM3_BURMA	Q62gm3 burkholderi
62	20	95.2	60	2	Q63Q29_BURPS	Q63q29 burkholderi
63	20	95.2	60	2	Q68UX3_9GEMI	Q68ux3 ageratum ye
64	20	95.2	61	2	Q8WYD0_HUMAN	Q8wyd0 homo sapien
65	20	95.2	61	2	Q82X74_NITEU	Q82x74 nitrosomona
66	20	95.2	61	2	Q98J39_RHILO	Q98j39 rhizobium l
67	20	95.2	61	2	Q9K743_BACHD	Q9k743 bacillus ha
68	20	95.2	62	2	Q5WHN0_BACSK	Q5whn0 bacillus cl
69	20	95.2	62	2	Q4GXV5_9GEMI	Q4gxv5 tomato gemi
70	20	95.2	63	2	Q5BSP2_SCHJA	Q5bsp2 schistosoma
71	20	95.2	63	2	Q663M1_YERPS	Q663m1 yersinia ps
72	20	95.2	63	2	Q9QCK2_BDV	Q9qck2 borna disea
73	20	95.2	63	2	Q9QCK5_BDV	Q9qck5 borna disea
74	20	95.2	63	2	Q9QCK8_BDV	Q9qck8 borna disea
75	20	95.2	63	2	Q9QCL1_BDV	Q9qcl1 borna disea
76	20	95.2	63	2	Q9QCL4_BDV	Q9qcl4 borna disea
77	20	95.2	63	2	Q9QCL7_BDV	Q9qcl7 borna disea
78	20	95.2	63	2	Q9QCM0_BDV	Q9qcm0 borna disea
79	20	95.2	63	2	Q9QCM3_BDV	Q9qcm3 borna disea
80	20	95.2	64	2	Q6ADC8_LEIXX	Q6adc8 leifsonia x
81	20	95.2	64	2	Q6ZX71_9GEMI	Q6zx71 papaya leaf
82	20	95.2	64	2	Q6ZX75_9GEMI	Q6zx75 papaya leaf
83	20	95.2	65	2	Q8PXM9_METMA	Q8pxm9 methanosarc
84	20	95.2	65	2	Q84KT4_BRAOL	Q84kt4 brassica ol
85	20	95.2	65	2	Q6ZX79_9GEMI	Q6zx79 tomato leaf
86	20	95.2	65	2	Q6ZX87_9GEMI	Q6zx87 tomato leaf
87	20	95.2	65	2	Q6ZX89_9GEMI	Q6zx89 ageratum ye
88	20	95.2	66	2	Q934F4_TREPH	Q934f4 treponema p
89	20	95.2	66	2	Q93DL5_TREPA	Q93dl5 treponema p
90	20	95.2	66	2	Q93DL6_9SPIO	Q93dl6 spirochete
91	20	95.2	66	2	Q93DL7_TREMD	Q93dl7 treponema m
92	20	95.2	66	2	Q93DL8_TRESO	Q93dl8 treponema s
93	20	95.2	66	2	Q93DL9_TREDE	Q93dl9 treponema d
94	20	95.2	66	2	Q93DM0_9SPIO	Q93dm0 spirochete
95	20	95.2	66	2	Q93DM1_9SPIO	Q93dm1 spirochete
96	20	95.2	66	2	Q93DM2_9SPIO	Q93dm2 spirochete
97	20	95.2	66	2	Q8ABX4_BACTN	Q8abx4 bacteroides
98	20	95.2	66	2	Q68UV4_9GEMI	Q68uv4 tomato leaf
99	20	95.2	66	2	Q6ZX73_9GEMI	Q6zx73 papaya leaf
100	20	95.2	66	2	Q6ZX85_9GEMI	Q6zx85 tomato leaf
101	20	95.2	67	2	Q6NS35_HUMAN	Q6ns35 homo sapien
102	20	95.2	67	2	Q5TMI2_ANOGA	Q5tmi2 anopheles g
103	20	95.2	67	2	Q41381_9ASTR	Q41381 senecio odo
104	20	95.2	67	2	Q6ZX83_9GEMI	Q6zx83 tomato leaf

105	20	95.2	68	2	Q757Q8_ASHGO	Q757q8	ashbya goss
106	20	95.2	68	2	Q4G0U0_HUMAN	Q4g0u0	homo sapien
107	20	95.2	68	2	Q7X6G5_ORYSA	Q7x6g5	oryza sativ
108	20	95.2	68	2	Q7F9H3_ORYSA	Q7f9h3	oryza sativ
109	20	95.2	68	2	Q6HFQ9_BACHK	Q6hfq9	bacillus th
110	20	95.2	68	2	O92602_TYLCV	O92602	tomato yell
111	20	95.2	68	2	Q6KCL1_9GEMI	Q6kcl1	tomato mayo
112	20	95.2	68	2	Q6KCL3_9GEMI	Q6kcl3	tomato mayo
113	20	95.2	68	2	Q6KCL5_9GEMI	Q6kcl5	tomato mayo
114	20	95.2	68	2	Q6ZX59_9GEMI	Q6zx59	papaya leaf
115	20	95.2	68	2	Q6ZX61_9GEMI	Q6zx61	papaya leaf
116	20	95.2	68	2	Q6ZX63_9GEMI	Q6zx63	papaya leaf
117	20	95.2	68	2	Q6ZX67_9GEMI	Q6zx67	papaya leaf
118	20	95.2	68	2	Q6ZX69_9GEMI	Q6zx69	papaya leaf
119	20	95.2	68	2	Q6ZX77_9GEMI	Q6zx77	squash leaf
120	20	95.2	68	2	Q6ZX81_9GEMI	Q6zx81	tomato leaf
121	20	95.2	68	2	Q8UZB8_9GEMI	Q8uzb8	tomato bego
122	20	95.2	68	2	Q8UZC0_9GEMI	Q8uzc0	tomato bego
123	20	95.2	68	2	Q9YKD4_TYLCV	Q9ykd4	tomato yell
124	20	95.2	69	2	Q6BGY5_DEBHA	Q6bgy5	debaryomyce
125	20	95.2	69	2	Q6ZX55_9GEMI	Q6zx55	papaya leaf
126	20	95.2	69	2	Q6ZX65_9GEMI	Q6zx65	papaya leaf
127	20	95.2	70	2	Q7WHS6_BORBR	Q7whs6	bordetella
128	20	95.2	70	2	Q5TJC8_9GEMI	Q5tjc8	papaya leaf
129	20	95.2	70	2	Q5TJD0_9GEMI	Q5tjd0	papaya leaf
130	20	95.2	70	2	Q6ZX57_9GEMI	Q6zx57	papaya leaf
131	20	95.2	70	2	Q8B8P3_9GEMI	Q8b8p3	tomato gemi
132	20	95.2	70	2	Q8B8P4_9GEMI	Q8b8p4	tomato gemi
133	20	95.2	70	2	Q8B8P6_9GEMI	Q8b8p6	tomato gemi
134	20	95.2	70	2	Q8B8P7_9GEMI	Q8b8p7	tomato gemi
135	20	95.2	70	2	Q8B8P8_9GEMI	Q8b8p8	tomato gemi
136	20	95.2	70	2	Q4ST10_TETNG	Q4st10	tetraodon n
137	20	95.2	71	1	TRAD_RHISN	P55420	rhizobium s
138	20	95.2	71	1	Y1435_METJA	Q58830	methanococc
139	20	95.2	71	2	Q8G9P7_ENTAG	Q8g9p7	enterobacte
140	20	95.2	71	2	Q5JZL7_9GEMI	Q5jzl7	papaya leaf
141	20	95.2	71	2	Q68UU8_9GEMI	Q68uu8	papaya leaf
142	20	95.2	72	2	Q7MBK8_VIBVY	Q7mbk8	vibrio vuln
143	20	95.2	72	2	Q9RI20_YERPE	Q9ri20	yersinia pe
144	20	95.2	72	2	Q5TJD2_9GEMI	Q5tjd2	papaya leaf
145	20	95.2	72	2	Q6R2T8_PASDO	Q6r2t8	passer dome
146	20	95.2	73	2	Q7Y2A6_9CAUD	Q7y2a6	phage phi 4
147	20	95.2	73	2	Q7Y2N3_9CAUD	Q7y2n3	stx2 conver
148	20	95.2	73	2	Q7Y344_9CAUD	Q7y344	stx1 conver
149	20	95.2	73	2	Q8SC56_9CAUD	Q8sc56	stx2 conver
150	20	95.2	73	2	Q8X3F3_ECO57	Q8x3f3	escherichia
151	20	95.2	73	2	Q9EYA1_ECO57	Q9eya1	escherichia
152	20	95.2	73	2	Q9KXG7_ECO57	Q9kxg7	escherichia
153	20	95.2	73	2	Q7TEL8_9GEMI	Q7tel8	tomato gemi
154	20	95.2	73	2	Q4ZIJ7_9PARA	Q4zij7	newcastle d
155	20	95.2	73	2	Q4ZIJ8_9PARA	Q4zij8	newcastle d
156	20	95.2	73	2	Q4ZIJ9_9PARA	Q4zij9	newcastle d
157	20	95.2	73	2	Q4ZIK0_9PARA	Q4zik0	newcastle d
158	20	95.2	73	2	Q4ZIK1_9PARA	Q4zik1	newcastle d
159	20	95.2	74	2	Q8R8F5_THETN	Q8r8f5	thermoanaer
160	20	95.2	74	2	Q9PGQ6_XYLFA	Q9pgq6	xylella fas
161	20	95.2	74	2	Q68UV2_9GEMI	Q68uv2	euphorbia l
162	20	95.2	74	2	Q7TGS1_9GEMI	Q7tgs1	malvastrum
163	20	95.2	74	2	Q8B6S8_TYLCV	Q8b6s8	tomato yell
164	20	95.2	74	2	Q8B6T0_TYLCV	Q8b6t0	tomato yell
165	20	95.2	74	2	Q8B6T2_TYLCV	Q8b6t2	tomato yell
166	20	95.2	74	2	Q8B6V7_9GEMI	Q8b6v7	tobacco lea
167	20	95.2	74	2	Q8B6V8_9GEMI	Q8b6v8	tobacco lea
168	20	95.2	74	2	Q8B6W0_9GEMI	Q8b6w0	tobacco lea
169	20	95.2	74	2	Q8B6W1_9GEMI	Q8b6w1	tobacco cur
170	20	95.2	74	2	Q8B6W2_9GEMI	Q8b6w2	tobacco cur
171	20	95.2	74	2	Q8B6W4_9GEMI	Q8b6w4	tobacco cur
172	20	95.2	74	2	Q8B6W6_9GEMI	Q8b6w6	tobacco cur
173	20	95.2	74	2	Q8B6W7_9GEMI	Q8b6w7	tobacco cur
174	20	95.2	74	2	Q8B6W8_9GEMI	Q8b6w8	tobacco cur
175	20	95.2	74	2	Q8B6X0_9GEMI	Q8b6x0	tobacco cur
176	20	95.2	74	2	Q8B6X1_9GEMI	Q8b6x1	tobacco cur
177	20	95.2	74	2	Q8B6X2_9GEMI	Q8b6x2	tobacco cur
178	20	95.2	74	2	Q8B8P9_9GEMI	Q8b8p9	tomato gemi
179	20	95.2	74	2	Q8B8Q0_9GEMI	Q8b8q0	tomato gemi
180	20	95.2	74	2	Q8B8Q1_9GEMI	Q8b8q1	tomato gemi
181	20	95.2	74	2	Q8B9F9_NPVRO	Q8b9f9	rachiplusia
182	20	95.2	74	2	Q8UYP5_9GEMI	Q8uyp5	tobacco lea
183	20	95.2	74	2	Q8UYP7_9GEMI	Q8uyp7	tobacco lea
184	20	95.2	74	2	Q4VYJ1_9GEMI	Q4vyj1	tomato yell
185	20	95.2	74	2	Q4VYJ3_9GEMI	Q4vyj3	tomato yell
186	20	95.2	74	2	Q4VYJ5_9GEMI	Q4vyj5	tomato yell
187	20	95.2	74	2	Q4VYJ7_9GEMI	Q4vyj7	tomato yell
188	20	95.2	74	2	Q4VYJ9_9GEMI	Q4vyj9	tomato yell
189	20	95.2	74	2	Q4VYK1_9GEMI	Q4vyk1	tomato yell
190	20	95.2	74	2	Q4VYK3_9GEMI	Q4vyk3	tomato yell
191	20	95.2	74	2	Q4VYK5_9GEMI	Q4vyk5	tomato yell
192	20	95.2	74	2	Q4VYK7_9GEMI	Q4vyk7	tomato yell
193	20	95.2	74	2	Q4VYK9_9GEMI	Q4vyk9	tomato yell
194	20	95.2	74	2	Q4VYL1_9GEMI	Q4vyl1	tobacco cur
195	20	95.2	74	2	Q4VYL3_9GEMI	Q4vyl3	tobacco cur
196	20	95.2	74	2	Q4VYL7_9GEMI	Q4vyl7	tobacco cur
197	20	95.2	74	2	Q4VYL9_9GEMI	Q4vyl9	tobacco lea
198	20	95.2	74	2	Q4VYM1_9GEMI	Q4vym1	tobacco lea
199	20	95.2	74	2	Q4VYM3_9GEMI	Q4vym3	tobacco lea
200	20	95.2	74	2	Q4VYM5_9GEMI	Q4vym5	tobacco lea
201	20	95.2	74	2	Q4VYM7_9GEMI	Q4vym7	tobacco lea
202	20	95.2	74	2	Q4VYM9_9GEMI	Q4vym9	malvastrum
203	20	95.2	74	2	Q4VYN1_9GEMI	Q4vyn1	malvastrum
204	20	95.2	74	2	Q4VYN3_9GEMI	Q4vyn3	malvastrum
205	20	95.2	74	2	Q4VYN5_9GEMI	Q4vyn5	malvastrum
206	20	95.2	74	2	Q4VYN7_9GEMI	Q4vyn7	malvastrum
207	20	95.2	74	2	Q4VYP1_9GEMI	Q4vyp1	malvastrum
208	20	95.2	74	2	Q4VYP3_9GEMI	Q4vyp3	malvastrum
209	20	95.2	74	2	Q4VYP5_9GEMI	Q4vyp5	malvastrum
210	20	95.2	74	2	Q4VYP7_9GEMI	Q4vyp7	malvastrum
211	20	95.2	74	2	Q4VYP9_9GEMI	Q4vyp9	malvastrum
212	20	95.2	74	2	Q4VYQ1_9GEMI	Q4vyq1	malvastrum
213	20	95.2	74	2	Q4VYQ5_9GEMI	Q4vyq5	tomato yell
214	20	95.2	74	2	Q4VYQ7_9GEMI	Q4vyq7	tomato yell
215	20	95.2	74	2	Q4VYQ9_9GEMI	Q4vyq9	tomato yell
216	20	95.2	74	2	Q4VYR3_9GEMI	Q4vyr3	tomato yell
217	20	95.2	74	2	Q4VYR5_9GEMI	Q4vyr5	tomato yell
218	20	95.2	74	2	Q4GXW1_9GEMI	Q4gxw1	tomato gemi
219	20	95.2	75	2	Q6Z8U5_ORYSA	Q6z8u5	oryza sativ
220	20	95.2	75	2	Q9S0S0_STRCO	Q9s0s0	streptomyce
221	20	95.2	75	2	Q9A0Q9_STRPY	Q9a0q9	streptococc
222	20	95.2	75	2	Q5WM52_9GEMI	Q5wm52	papaya leaf
223	20	95.2	75	2	Q5WM54_9GEMI	Q5wm54	papaya leaf
224	20	95.2	75	2	Q5WM70_9GEMI	Q5wm70	ageratum ye
225	20	95.2	75	2	Q5WM72_9GEMI	Q5wm72	ageratum ye
226	20	95.2	75	2	Q5WM74_9GEMI	Q5wm74	ageratum ye
227	20	95.2	75	2	Q5WM76_9GEMI	Q5wm76	ageratum ye
228	20	95.2	75	2	Q5WM78_9GEMI	Q5wm78	ageratum ye
229	20	95.2	75	2	Q5WM80_9GEMI	Q5wm80	ageratum ye
230	20	95.2	75	2	Q5WM82_9GEMI	Q5wm82	ageratum ye
231	20	95.2	75	2	Q5WM84_9GEMI	Q5wm84	ageratum ye
232	20	95.2	75	2	Q5WM86_9GEMI	Q5wm86	ageratum ye
233	20	95.2	75	2	Q5WM88_9GEMI	Q5wm88	ageratum ye
234	20	95.2	75	2	Q5WM90_9GEMI	Q5wm90	ageratum ye
235	20	95.2	75	2	Q5WM92_9GEMI	Q5wm92	ageratum ye
236	20	95.2	75	2	Q5WM94_9GEMI	Q5wm94	ageratum ye
237	20	95.2	75	2	Q5WM98_9GEMI	Q5wm98	ageratum ye
238	20	95.2	75	2	Q5WMA0_9GEMI	Q5wma0	ageratum ye
239	20	95.2	75	2	Q6A0F5_9GEMI	Q6a0f5	stachytarph
240	20	95.2	75	2	Q6A0F7_9GEMI	Q6a0f7	stachytarph
241	20	95.2	75	2	Q6A0F9_9GEMI	Q6a0f9	stachytarph
242	20	95.2	75	2	Q6A0G1_9GEMI	Q6a0g1	stachytarph
243	20	95.2	75	2	Q6A0G3_9GEMI	Q6a0g3	stachytarph
244	20	95.2	75	2	Q6A0G5_9GEMI	Q6a0g5	stachytarph
245	20	95.2	75	2	Q6A0G7_9GEMI	Q6a0g7	stachytarph
246	20	95.2	75	2	Q6A0G9_9GEMI	Q6a0g9	stachytarph
247	20	95.2	75	2	Q7TGS2_9GEMI	Q7tgs2	malvastrum
248	20	95.2	75	2	Q8B8R1_9GEMI	Q8b8r1	ageratum ye
249	20	95.2	75	2	Q8B8R3_9GEMI	Q8b8r3	ageratum ye
250	20	95.2	75	2	Q9YRC9_9PARA	Q9yrc9	newcastle d

251	20	95.2	75	2	Q68V32_9GEMI	Q68v32	sida yellow
252	20	95.2	75	2	Q68V43_9GEMI	Q68v43	sida yellow
253	20	95.2	75	2	Q4VYL5_9GEMI	Q4vyl5	tobacco cur
254	20	95.2	75	2	Q4VZ16_9GEMI	Q4vz16	papaya leaf
255	20	95.2	75	2	Q4VZ18_9GEMI	Q4vz18	papaya leaf
256	20	95.2	75	2	Q4VZ20_9GEMI	Q4vz20	ageratum ye
257	20	95.2	75	2	Q4VZ23_9GEMI	Q4vz23	ageratum ye
258	20	95.2	75	2	Q4U0F5_9GEMI	Q4u0f5	hibiscus ro
259	20	95.2	75	2	Q4U0F7_9GEMI	Q4u0f7	lindernia a
260	20	95.2	75	2	Q4U0F9_9GEMI	Q4u0f9	lindernia a
261	20	95.2	75	2	Q5JZK9_9GEMI	Q5jzk9	senecio yel
262	20	95.2	75	2	Q5JZK7_9GEMI	Q5jzk7	senecio yel
263	20	95.2	76	2	Q79EL6_ECOLI	Q79el6	escherichia
264	20	95.2	76	2	P76024_ECOLI	P76024	escherichia
265	20	95.2	76	2	Q5WM96_9GEMI	Q5wm96	ageratum ye
266	20	95.2	77	2	Q64DT2_9ARCH	Q64dt2	uncultured
267	20	95.2	77	2	Q5TUC9_ANOGA	Q5tuc9	anopheles g
268	20	95.2	77	2	Q50SR4_ENTHI	Q50sr4	entamoeba h
269	20	95.2	77	2	O36988_9PARA	O36988	newcastle d
270	20	95.2	77	2	O36990_9PARA	O36990	newcastle d
271	20	95.2	77	2	O36992_9PARA	O36992	newcastle d
272	20	95.2	77	2	O36994_9PARA	O36994	newcastle d
273	20	95.2	77	2	O36996_9PARA	O36996	newcastle d
274	20	95.2	77	2	O36998_9PARA	O36998	newcastle d
275	20	95.2	77	2	O37000_9PARA	O37000	newcastle d
276	20	95.2	77	2	O37002_9PARA	O37002	newcastle d
277	20	95.2	77	2	O37004_9PARA	O37004	newcastle d
278	20	95.2	77	2	O37006_9PARA	O37006	newcastle d
279	20	95.2	77	2	O37008_9PARA	O37008	newcastle d
280	20	95.2	77	2	O37010_9PARA	O37010	newcastle d
281	20	95.2	77	2	O37012_9PARA	O37012	newcastle d
282	20	95.2	77	2	O37014_9PARA	O37014	newcastle d
283	20	95.2	77	2	Q68UY3_9GEMI	Q68uy3	papaya leaf
284	20	95.2	77	2	Q6SEG6_9PARA	Q6seg6	newcastle d
285	20	95.2	77	2	Q6SEG7_9PARA	Q6seg7	newcastle d
286	20	95.2	77	2	Q6SEG8_9PARA	Q6seg8	newcastle d
287	20	95.2	77	2	Q6SEG9_9PARA	Q6seg9	newcastle d
288	20	95.2	77	2	Q83780_9PARA	Q83780	newcastle d
289	20	95.2	77	2	Q83782_9PARA	Q83782	newcastle d
290	20	95.2	77	2	Q83790_9PARA	Q83790	newcastle d
291	20	95.2	77	2	Q83800_9PARA	Q83800	newcastle d
292	20	95.2	77	2	Q83820_9PARA	Q83820	newcastle d
293	20	95.2	77	2	Q83824_9PARA	Q83824	newcastle d
294	20	95.2	77	2	Q9DKI4_9PARA	Q9dki4	newcastle d
295	20	95.2	77	2	Q9DKI5_9PARA	Q9dki5	newcastle d
296	20	95.2	77	2	Q9DKI6_9PARA	Q9dki6	newcastle d
297	20	95.2	77	2	Q9DKI7_9PARA	Q9dki7	newcastle d
298	20	95.2	77	2	Q9DKI8_9PARA	Q9dki8	newcastle d
299	20	95.2	77	2	Q9DKI9_9PARA	Q9dki9	newcastle d
300	20	95.2	77	2	Q9DKJ0_9PARA	Q9dkj0	newcastle d
301	20	95.2	77	2	Q9J3V8_9PARA	Q9j3v8	newcastle d
302	20	95.2	77	2	Q9J3W0_9PARA	Q9j3w0	newcastle d
303	20	95.2	77	2	Q9J3W4_9PARA	Q9j3w4	newcastle d
304	20	95.2	77	2	Q9J3X0_9PARA	Q9j3x0	newcastle d
305	20	95.2	77	2	Q9J3X4_9PARA	Q9j3x4	newcastle d
306	20	95.2	77	2	Q9J3X6_9PARA	Q9j3x6	newcastle d
307	20	95.2	77	2	Q9J3X8_9PARA	Q9j3x8	newcastle d
308	20	95.2	77	2	Q9J3Y0_9PARA	Q9j3y0	newcastle d
309	20	95.2	77	2	Q9QPC7_TYLCV	Q9qpc7	tomato yell
310	20	95.2	78	2	Q972N2_SULTO	Q972n2	sulfolobus
311	20	95.2	78	2	Q7PL20_ANOGA	Q7pl20	anopheles g
312	20	95.2	78	2	Q5UKX2_9BROM	Q5ukx2	fragaria ch
313	20	95.2	79	2	O14891_HUMAN	O14891	homo sapien
314	20	95.2	79	2	Q6P2M4_HUMAN	Q6p2m4	homo sapien
315	20	95.2	79	2	Q4VYN9_9GEMI	Q4vyn9	malvastrum
316	20	95.2	79	2	Q4VYQ3_9GEMI	Q4vyq3	malvastrum
317	20	95.2	79	2	Q9WPL8_9GEMI	Q9wpl8	tomato leaf
318	20	95.2	80	2	Q8VWT5_NARPS	Q8vwt5	narcissus p
319	20	95.2	80	2	Q57TK5_SALCH	Q57tk5	salmonella
320	20	95.2	80	2	Q8X2N0_ECO57	Q8x2n0	escherichia
321	20	95.2	80	2	Q5WGQ0_BACSK	Q5wgq0	bacillus cl
322	20	95.2	80	2	Q82UA4_NITEU	Q82ua4	nitrosomona
323	20	95.2	80	2	Q839M8_ENTFA	Q839m8	enterococcu
324	20	95.2	81	1	OADG2_SALTY	P58650	salmonella
325	20	95.2	82	2	Q8GWY6_ARATH	Q8gwy6	arabidopsis
326	20	95.2	82	2	Q4KM25_RAT	Q4km25	rattus norv
327	20	95.2	83	2	Q54SS4_DICDI	Q54ss4	dictyosteli
328	20	95.2	83	2	Q4X4S6_PLACH	Q4x4s6	plasmodium
329	20	95.2	83	2	Q6K6S3_ORYSA	Q6k6s3	oryza sativ
330	20	95.2	83	2	Q6MXB9_SERMA	Q6mxb9	serratia ma
331	20	95.2	83	2	Q847N6_ASTYP	Q847n6	aster yello
332	20	95.2	83	2	Q4IVY5_AZOVI	Q4ivy5	azotobacter
333	20	95.2	83	2	Q7W6U0_BORPA	Q7w6u0	bordetella
334	20	95.2	85	2	Q9N114_SHEEP	Q9n114	ovis aries
335	20	95.2	85	2	Q96352_BRANA	Q96352	brassica na
336	20	95.2	85	2	Q5M0C8_STRT1	Q5m0c8	streptococc
337	20	95.2	85	2	Q5M4X5_STRT2	Q5m4x5	streptococc
338	20	95.2	86	2	Q570C7_ARATH	Q570c7	arabidopsis
339	20	95.2	86	2	Q4LXY7_9BURK	Q4lxy7	burkholderi
340	20	95.2	86	2	Q82UX2_NITEU	Q82ux2	nitrosomona
341	20	95.2	86	2	Q7UVE5_RHOBA	Q7uve5	rhodopirell
342	20	95.2	87	1	SSS1_SCYCA	P13275	scyliorhinu
343	20	95.2	87	2	Q5CKC6_CRYHO	Q5ckc6	cryptospori
344	20	95.2	87	2	Q94MP0_9CAUD	Q94mp0	bacteriopha
345	20	95.2	87	2	Q5GL76_BDV	Q5gl76	borna disea
346	20	95.2	87	2	Q5GL97_BDV	Q5gl97	borna disea
347	20	95.2	87	2	Q5GLA6_BDV	Q5gla6	borna disea
348	20	95.2	87	2	Q5GLD0_BDV	Q5gld0	borna disea
349	20	95.2	87	2	Q86622_BDV	Q86622	borna disea
350	20	95.2	87	2	Q8JJK0_BDV	Q8jjk0	borna disea
351	20	95.2	87	2	Q912Z2_BDV	Q912z2	borna disea
352	20	95.2	87	2	Q912Z5_BDV	Q912z5	borna disea
353	20	95.2	87	2	Q912Z9_BDV	Q912z9	borna disea
354	20	95.2	87	2	Q9Q9V0_BDV	Q9q9v0	borna disea
355	20	95.2	87	2	Q9WNA1_BDV	Q9wna1	borna disea
356	20	95.2	88	1	Y778_STRA5	Q8e0f5	streptococc
357	20	95.2	88	1	Y798_STRA3	Q8e629	streptococc
358	20	95.2	89	1	VP54_BPAPS	Q9t1p4	bacteriopha
359	20	95.2	89	2	Q6H0Y4_9CREN	Q6h0y4	sulfolobus
360	20	95.2	89	2	Q97WJ2_SULSO	Q97wj2	sulfolobus
361	20	95.2	89	2	Q67TC5_SYMTH	Q67tc5	symbiobacte
362	20	95.2	90	1	FETP_PSEAE	Q9hu36	pseudomonas
363	20	95.2	90	1	FETP_PSESM	Q87uf5	pseudomonas
364	20	95.2	90	2	Q5W667_ORYSA	Q5w667	oryza sativ
365	20	95.2	90	2	Q4ZLP3_PSESY	Q4zlp3	pseudomonas
366	20	95.2	90	2	Q4J228_AZOVI	Q4j228	azotobacter
367	20	95.2	91	2	Q5TXS7_ANOGA	Q5txs7	anopheles g
368	20	95.2	91	2	Q6LW69_PHOPR	Q6lw69	photobacter
369	20	95.2	91	2	Q8D665_VIBVU	Q8d665	vibrio vuln
370	20	95.2	91	2	Q8R9Q8_THETN	Q8r9q8	thermoanaer
371	20	95.2	91	2	Q92XL9_RHIME	Q92xl9	rhizobium m
372	20	95.2	91	2	Q7UM06_RHOBA	Q7um06	rhodopirell
373	20	95.2	91	2	Q53CW8_9GAMA	Q53cw8	macaca fusc
374	20	95.2	92	2	Q8ZB80_YERPE	Q8zb80	yersinia pe
375	20	95.2	92	2	Q66FA3_YERPS	Q66fa3	yersinia ps
376	20	95.2	92	2	Q82GZ5_STRAW	Q82gz5	streptomyce
377	20	95.2	92	2	Q8KBJ4_CHLTE	Q8kbj4	chlorobium
378	20	95.2	93	2	Q5TYE6_ANOGA	Q5tye6	anopheles g
379	20	95.2	93	2	Q9BLC2_CAEEL	Q9blc2	caenorhabdi
380	20	95.2	93	2	Q4QYB7_9ACTO	Q4qyb7	streptomyce
381	20	95.2	93	2	Q4QYC4_STRLI	Q4qyc4	streptomyce
382	20	95.2	93	2	Q4QYC5_9ACTO	Q4qyc5	streptomyce
383	20	95.2	93	2	Q4QYC6_STRAM	Q4qyc6	streptomyce
384	20	95.2	94	1	ABC1_BPP22	P11190	bacteriopha
385	20	95.2	94	2	Q9VQ55_DROME	Q9vq55	drosophila
386	20	95.2	94	2	Q76H41_9CAUD	Q76h41	salmonella
387	20	95.2	94	2	Q8SCE8_9VIRU	Q8sce8	vibrio harv
388	20	95.2	94	2	Q8HAH6_BPST6	Q8hah6	bacteriopha
389	20	95.2	94	2	Q77D70_BPP22	Q77d70	bacteriopha
390	20	95.2	94	2	Q5PF59_SALPA	Q5pf59	salmonella
391	20	95.2	94	2	Q8E1D3_STRA5	Q8e1d3	streptococc
392	20	95.2	94	2	Q8E6V0_STRA3	Q8e6v0	streptococc
393	20	95.2	95	2	Q7WZT3_BORGA	Q7wzt3	borrelia ga
394	20	95.2	95	2	Q6LHW7_PHOPR	Q6lhw7	photobacter
395	20	95.2	95	2	Q8ZF39_YERPE	Q8zf39	yersinia pe
396	20	95.2	95	2	Q66B78_YERPS	Q66b78	yersinia ps

397	20	95.2	95	2	Q62B28_BURMA	Q62b28	burkholderi
398	20	95.2	96	1	EXOX_RHISN	P14185	rhizobium s
399	20	95.2	96	1	PAPA5_CARPA	P05993	carica papa
400	20	95.2	96	2	Q50ZD8_ENTHI	Q50zd8	entamoeba h
401	20	95.2	96	2	Q8W6V6_9CAUD	Q8w6v6	cyanophage
402	20	95.2	96	2	Q7WZT4_BORAF	Q7wzt4	borrelia af
403	20	95.2	96	2	Q6W2H5_RHISN	Q6w2h5	rhizobium s
404	20	95.2	96	2	Q5FNW8_GLUOX	Q5fnw8	gluconobact
405	20	95.2	96	2	Q6AJV7_DESPS	Q6ajv7	desulfotale
406	20	95.2	97	2	Q7WZT2_9SPIO	Q7wzt2	borrelia va
407	20	95.2	97	2	Q6LWA7_PHOPR	Q6lwa7	photobacter
408	20	95.2	97	2	Q8D647_VIBVU	Q8d647	vibrio vuln
409	20	95.2	98	1	EXOX_RHIME	Q02730	rhizobium m
410	20	95.2	98	2	Q24074_DROME	Q24074	drosophila
411	20	95.2	98	2	Q9VUS9_DROME	Q9vus9	drosophila
412	20	95.2	98	2	Q50R98_ENTHI	Q50r98	entamoeba h
413	20	95.2	98	2	Q6EPF9_ORYSA	Q6epf9	oryza sativ
414	20	95.2	98	2	Q8EPC8_OCEIH	Q8epc8	oceanobacil
415	20	95.2	98	2	Q92TK8_RHIME	Q92tk8	rhizobium m
416	20	95.2	99	2	Q4Y8V2_PLACH	Q4y8v2	plasmodium
417	20	95.2	99	2	Q9L5I5_SALTI	Q9l5i5	salmonella
418	20	95.2	99	2	Q9Q0B0_NPVAG	Q9q0b0	anticarsia
419	20	95.2	99	2	Q9QMX1_9VIRU	Q9qmx1	hantaan vir
420	20	95.2	100	1	URE3_PROMM	Q7v3v4	prochloroco
421	20	95.2	100	1	URE3_PROMP	Q7v1b4	prochloroco
422	20	95.2	100	1	URE3_PROS9	Q91642	prochloroco
423	20	95.2	100	1	URE3_SYNPV	O87400	synechococc
424	20	95.2	100	1	URE3_SYNPX	Q7u3i5	synechococc
425	20	95.2	100	2	Q5AXW8_EMENI	Q5axw8	aspergillus
426	20	95.2	100	2	Q27316_DROME	Q27316	drosophila
427	20	95.2	100	2	Q9VUS3_DROME	Q9vus3	drosophila
428	20	95.2	100	2	Q5IGZ4_9CAUD	Q5igz4	bacteriopha
429	20	95.2	101	2	Q8I7Q4_DROME	Q8i7q4	drosophila
430	20	95.2	101	2	Q8RZU9_ORYSA	Q8rzu9	oryza sativ
431	20	95.2	101	2	Q6WKW5_9ACTO	Q6wkw5	streptomyce
432	20	95.2	101	2	Q6WL11_9ACTO	Q6wl11	streptomyce
433	20	95.2	101	2	Q6WL18_9ACTO	Q6wl18	streptomyce
434	20	95.2	101	2	Q6WL38_STRST	Q6wl38	streptomyce
435	20	95.2	101	2	Q6WL44_STRPH	Q6wl44	streptomyce
436	20	95.2	101	2	Q6WL62_STRCO	Q6wl62	streptomyce
437	20	95.2	101	2	Q6WL69_STRAV	Q6wl69	streptomyce
438	20	95.2	101	2	Q6WL73_9ACTO	Q6wl73	streptomyce
439	20	95.2	101	2	Q6WL78_STRAO	Q6wl78	streptomyce
440	20	95.2	101	2	Q9RR17_MYCCE	Q9rr17	mycobacteri
441	20	95.2	101	2	Q6QMZ5_CHILA	Q6qmz5	chinchilla
442	20	95.2	101	2	Q91GI2_NPVEP	Q91gi2	epiphyas po
443	20	95.2	102	2	Q59X30_CANAL	Q59x30	candida alb
444	20	95.2	102	2	Q7S5G0_NEUCR	Q7s5g0	neurospora
445	20	95.2	102	2	Q70IL6_9PSED	Q70il6	pseudomonas
446	20	95.2	102	2	Q88HT0_PSEPK	Q88ht0	pseudomonas
447	20	95.2	102	2	Q82Q80_STRAW	Q82q80	streptomyce
448	20	95.2	103	2	Q9KLM2_VIBCH	Q9klm2	vibrio chol
449	20	95.2	103	2	Q8BQK2_MOUSE	Q8bqk2	mus musculu
450	20	95.2	104	2	Q9U5Z1_CAEEL	Q9u5z1	caenorhabdi
451	20	95.2	104	2	Q5GXN0_XANOR	Q5gxn0	xanthomonas
452	20	95.2	104	2	Q4K3L3_PSEF5	Q4k3l3	pseudomonas
453	20	95.2	104	2	Q8A641_BACTN	Q8a641	bacteroides
454	20	95.2	104	2	Q7ZZU0_OREMO	Q7zzu0	oreochromis
455	20	95.2	105	2	O24309_PEA	O24309	pisum sativ
456	20	95.2	105	2	Q88UH4_LACPL	Q88uh4	lactobacill
457	20	95.2	106	1	RL24_PORGI	Q7mtm4	porphyromon
458	20	95.2	106	2	Q6CKM8_KLULA	Q6ckm8	kluyveromyc
459	20	95.2	106	2	Q6NP30_DROME	Q6np30	drosophila
460	20	95.2	106	2	Q9MYZ6_TRIVU	Q9myz6	trichosurus
461	20	95.2	106	2	Q83EN0_COXBU	Q83en0	coxiella bu
462	20	95.2	107	2	Q58HJ1_SHEEP	Q58hj1	ovis aries
463	20	95.2	107	2	Q6WRV8_HAFAL	Q6wrv8	hafnia alve
464	20	95.2	107	2	Q6WRW5_HAFAL	Q6wrw5	hafnia alve
465	20	95.2	107	2	Q4FUK4_9GAMM	Q4fuk4	psychrobact
466	20	95.2	108	2	Q4WI77_ASPFU	Q4wi77	aspergillus
467	20	95.2	108	2	Q6Z2K8_ORYSA	Q6z2k8	oryza sativ
468	20	95.2	108	2	Q9ZU61_ARATH	Q9zu61	arabidopsis
469	20	95.2	108	2	Q81YG0_BACAN	Q81yg0	bacillus an
470	20	95.2	108	2	Q88U03_LACPL	Q88u03	lactobacill
471	20	95.2	108	2	Q60516_9MURI	Q60516	rattus sp.
472	20	95.2	108	2	Q4S4B7_TETNG	Q4s4b7	tetraodon n
473	20	95.2	109	2	Q50UB4_ENTHI	Q50ub4	entamoeba h
474	20	95.2	109	2	Q9ETL9_BORGA	Q9etl9	borrelia ga
475	20	95.2	109	2	Q9ETM0_BORAF	Q9etm0	borrelia af
476	20	95.2	109	2	Q9ETM1_BORBU	Q9etm1	borrelia bu
477	20	95.2	109	2	Q9F9V8_BORTU	Q9f9v8	borrelia tu
478	20	95.2	109	2	Q9F9V9_BORHE	Q9f9v9	borrelia he
479	20	95.2	109	2	Q9F9W0_BORAD	Q9f9w0	borrelia an
480	20	95.2	109	2	Q9F9W1_9SPIO	Q9f9w1	borrelia bi
481	20	95.2	109	2	Q9F9W2_9SPIO	Q9f9w2	borrelia lu
482	20	95.2	109	2	Q9F9W3_BORJA	Q9f9w3	borrelia ja
483	20	95.2	109	2	Q9F9W4_9SPIO	Q9f9w4	borrelia va
484	20	95.2	109	2	Q5XIS4_RAT	Q5xis4	rattus norv
485	20	95.2	110	2	O26682_METTH	O26682	methanobact
486	20	95.2	110	2	Q59NH4_CANAL	Q59nh4	candida alb
487	20	95.2	110	2	Q9C1L6_NEUCR	Q9c1l6	neurospora
488	20	95.2	110	2	Q6YSK3_BRANA	Q6ysk3	brassica na
489	20	95.2	110	2	Q6RGR6_STRLA	Q6rgr6	streptomyce
490	20	95.2	110	2	Q8XYK3_RALSO	Q8xyk3	ralstonia s
491	20	95.2	110	2	Q8ZDT7_YERPE	Q8zdt7	yersinia pe
492	20	95.2	110	2	Q8C5D2_MOUSE	Q8c5d2	mus musculu
493	20	95.2	110	2	Q5ZP61_9VIRU	Q5zp61	cotesia con
494	20	95.2	111	2	Q6NSH7_HUMAN	Q6nsh7	homo sapien
495	20	95.2	111	2	Q5ZBI6_ORYSA	Q5zbi6	oryza sativ
496	20	95.2	111	2	Q5ZDW6_ORYSA	Q5zdw6	oryza sativ
497	20	95.2	111	2	Q7WY00_PSEAE	Q7wy00	pseudomonas
498	20	95.2	111	2	Q5HLT5_STAEQ	Q5hlt5	staphylococ
499	20	95.2	111	2	Q89CV9_BRAJA	Q89cv9	bradyrhizob
500	20	95.2	111	2	Q877N3_PSEPK	Q877n3	pseudomonas
501	20	95.2	111	2	Q877M6_PSEPK	Q877m6	pseudomonas
502	20	95.2	112	2	O58145_PYRHO	O58145	pyrococcus
503	20	95.2	112	2	Q6BMJ2_DEBHA	Q6bmj2	debaryomyce
504	20	95.2	112	2	Q7QU11_GIALA	Q7qu11	giardia lam
505	20	95.2	112	2	Q7PUZ5_ANOGA	Q7puz5	anopheles g
506	20	95.2	112	2	Q69TP2_ORYSA	Q69tp2	oryza sativ
507	20	95.2	112	2	Q9FIG4_ARATH	Q9fig4	arabidopsis
508	20	95.2	112	2	Q84RD1_ARATH	Q84rd1	arabidopsis
509	20	95.2	112	2	Q4V2X5_BURMA	Q4v2x5	burkholderi
510	20	95.2	112	2	Q91BD0_NPVST	Q91bd0	spodoptera
511	20	95.2	113	1	RL34_METKA	Q8tzb2	methanopyru
512	20	95.2	113	2	Q50T98_ENTHI	Q50t98	entamoeba h
513	20	95.2	113	2	Q5YBC4_9CHLO	Q5ybc4	helicospori
514	20	95.2	113	2	Q94EX5_CUCME	Q94ex5	cucumis mel
515	20	95.2	113	2	Q9XCB3_RHOMR	Q9xcb3	rhodothermu
516	20	95.2	113	2	Q4J3D5_AZOVI	Q4j3d5	azotobacter
517	20	95.2	113	2	Q8UCV9_AGRT5	Q8ucv9	agrobacteri
518	20	95.2	113	2	Q9HVK9_PSEAE	Q9hvk9	pseudomonas
519	20	95.2	114	2	Q50ZF8_ENTHI	Q50zf8	entamoeba h
520	20	95.2	114	2	Q7BF85_YERPS	Q7bf85	yersinia ps
521	20	95.2	114	2	Q7VWS3_BORPE	Q7vws3	bordetella
522	20	95.2	114	2	Q7WJS4_BORBR	Q7wjs4	bordetella
523	20	95.2	114	2	Q93PD5_YERPE	Q93pd5	yersinia pe
524	20	95.2	114	2	Q66A07_YERPS	Q66a07	yersinia ps
525	20	95.2	114	2	Q8NPD4_CORGL	Q8npd4	corynebacte
526	20	95.2	114	2	Q8KDU0_CHLTE	Q8kdu0	chlorobium
527	20	95.2	114	2	Q8DG42_VIBVU	Q8dg42	vibrio vuln
528	20	95.2	114	2	Q6TPJ7_RAT	Q6tpj7	rattus norv
529	20	95.2	115	1	YG3Z_YEAST	P32475	saccharomyc
530	20	95.2	115	2	Q518F0_ENTHI	Q518f0	entamoeba h
531	20	95.2	115	2	Q4V477_DROME	Q4v477	drosophila
532	20	95.2	115	2	Q4YH95_PLABE	Q4yh95	plasmodium
533	20	95.2	115	2	Q4FUP3_9GAMM	Q4fup3	psychrobact
534	20	95.2	115	2	Q8ZIR6_YERPE	Q8zir6	yersinia pe
535	20	95.2	115	2	Q66EX3_YERPS	Q66ex3	yersinia ps
536	20	95.2	115	2	Q62BN5_BURMA	Q62bn5	burkholderi
537	20	95.2	115	2	Q8NPC0_CORGL	Q8npc0	corynebacte
538	20	95.2	115	2	Q8CRE1_STAEP	Q8cre1	staphylococ
539	20	95.2	116	2	Q56GD8_PEPCA	Q56gd8	peperomia c
540	20	95.2	116	2	Q6D9A9_ERWCT	Q6d9a9	erwinia car
541	20	95.2	116	2	Q80JM3_9RHAB	Q80jm3	rabies viru
542	20	95.2	116	2	Q80JM4_9RHAB	Q80jm4	rabies viru

543	20	95.2	116	2	Q80JM6_9RHAB	Q80jm6 rabies viru
544	20	95.2	116	2	Q80JM7_9RHAB	Q80jm7 rabies viru
545	20	95.2	116	2	Q80JM8_9RHAB	Q80jm8 rabies viru
546	20	95.2	116	2	Q80JM9_9RHAB	Q80jm9 rabies viru
547	20	95.2	116	2	Q80JN0_9RHAB	Q80jn0 rabies viru
548	20	95.2	116	2	Q80JN1_9RHAB	Q80jn1 rabies viru
549	20	95.2	116	2	Q80JN2_9RHAB	Q80jn2 rabies viru
550	20	95.2	116	2	Q80JN3_9RHAB	Q80jn3 rabies viru
551	20	95.2	116	2	Q80JN4_9RHAB	Q80jn4 rabies viru
552	20	95.2	116	2	Q80JN5_9RHAB	Q80jn5 rabies viru
553	20	95.2	116	2	Q80JN6_9RHAB	Q80jn6 rabies viru
554	20	95.2	116	2	O41053_CHVP1	O41053 paramecium
555	20	95.2	117	2	Q9Y6H7_HUMAN	Q9y6h7 homo sapien
556	20	95.2	117	2	Q4N0X1_THEPA	Q4n0x1 theileria p
557	20	95.2	117	2	Q7F0A4_ORYSA	Q7f0a4 oryza sativ
558	20	95.2	117	2	Q941S4_ORYSA	Q941s4 oryza sativ
559	20	95.2	117	2	Q56GH4_PHYPA	Q56gh4 physcomitre
560	20	95.2	117	2	Q8EL81_OCEIH	Q8el81 oceanobacil
561	20	95.2	118	1	ACPS_STRPY	Q99y97 streptococc
562	20	95.2	118	2	Q7CX77_AGRT5	Q7cx77 agrobacteri
563	20	95.2	118	2	Q5LNT4_SILPO	Q5lnt4 silicibacte
564	20	95.2	118	2	Q88UL4_LACPL	Q88ul4 lactobacill
565	20	95.2	118	2	Q4SBM5_TETNG	Q4sbm5 tetraodon n
566	20	95.2	119	2	Q8LR81_ORYSA	Q8lr81 oryza sativ
567	20	95.2	119	2	Q5CCY7_9GEMI	Q5ccy7 tomato leaf
568	20	95.2	119	2	Q4W4Y3_9GEMI	Q4w4y3 tomato leaf
569	20	95.2	120	2	Q9M714_9SOLA	Q9m714 petunia axi
570	20	95.2	120	2	Q8CL76_YERPE	Q8cl76 yersinia pe
571	20	95.2	120	2	Q5LEI4_BACFN	Q5lei4 bacteroides
572	20	95.2	120	2	Q6P6E5_BRARE	Q6p6e5 brachydanio
573	20	95.2	121	1	YXEE_BACSU	P54944 bacillus su
574	20	95.2	121	2	Q9YE40_AERPE	Q9ye40 aeropyrum p
575	20	95.2	121	2	Q4ZVD1_PSESY	Q4zvd1 pseudomonas
576	20	95.2	121	2	Q4ZYG7_PSESY	Q4zyg7 pseudomonas
577	20	95.2	121	2	Q6CYR3_ERWCT	Q6cyr3 erwinia car
578	20	95.2	121	2	Q5CCW1_9GEMI	Q5ccw1 ageratum ye
579	20	95.2	121	2	Q5CCW5_9GEMI	Q5ccw5 ageratum ye
580	20	95.2	121	2	Q5CCW9_9GEMI	Q5ccw9 ageratum ye
581	20	95.2	121	2	Q5CCX3_9GEMI	Q5ccx3 ageratum ye
582	20	95.2	121	2	Q5CCX7_9GEMI	Q5ccx7 ageratum ye
583	20	95.2	121	2	Q5CCZ5_9GEMI	Q5ccz5 ageratum ye
584	20	95.2	121	2	Q5DW23_9GEMI	Q5dw23 ageratum ye
585	20	95.2	122	1	M540_ARATH	P93309 arabidopsis
586	20	95.2	122	1	RL35_CHICK	Q98tf7 gallus gall
587	20	95.2	122	1	RL35_HIPCM	Q6uzf7 hippocampus
588	20	95.2	122	1	RL35_HUMAN	P42766 homo sapien
589	20	95.2	122	1	RL35_OPHHA	Q69cj9 ophiophagus
590	20	95.2	122	1	RL35_PIG	Q29361 sus scrofa
591	20	95.2	122	1	RL35_PLAFE	Q5dvh6 platichthys
592	20	95.2	122	1	RL35_XENTR	Q6pbc1 xenopus tro
593	20	95.2	122	2	Q8SVM9_ENCCU	Q8svm9 encephalito
594	20	95.2	122	2	Q6DR86_ARATH	Q6dr86 arabidopsis
595	20	95.2	122	2	Q4FR01_9GAMM	Q4fr01 psychrobact
596	20	95.2	122	2	Q7UM88_RHOBA	Q7um88 rhodopirell
597	20	95.2	122	2	Q74CF4_GEOSL	Q74cf4 geobacter s
598	20	95.2	122	2	Q5CCY3_9GEMI	Q5ccy3 pepper yell
599	20	95.2	122	2	Q5CCZ9_9GEMI	Q5ccz9 pepper yell
600	20	95.2	122	2	Q5PPY9_XENLA	Q5ppy9 xenopus lae
601	20	95.2	123	2	Q4WWK4_ASPFU	Q4wwk4 aspergillus
602	20	95.2	123	2	Q8N8S8_HUMAN	Q8n8s8 homo sapien
603	20	95.2	123	2	Q4VBY5_HUMAN	Q4vby5 homo sapien
604	20	95.2	123	2	Q4ZSN0_PSESY	Q4zsn0 pseudomonas
605	20	95.2	123	2	Q92ZF6_RHIME	Q92zf6 rhizobium m
606	20	95.2	123	2	Q5CCZ1_9GEMI	Q5ccz1 tomato leaf
607	20	95.2	123	2	Q4TAF6_TETNG	Q4taf6 tetraodon n
608	20	95.2	123	2	Q4FZQ7_XENLA	Q4fzq7 xenopus lae
609	20	95.2	124	2	Q8WZJ6_SCHPO	Q8wzj6 schizosacch
610	20	95.2	124	2	Q8VLY5_THIFE	Q8vly5 thiobacillu
611	20	95.2	124	2	Q5LBE8_BACFN	Q5lbe8 bacteroides
612	20	95.2	124	2	Q64RU6_BACFR	Q64ru6 bacteroides
613	20	95.2	124	2	Q4SGM6_TETNG	Q4sgm6 tetraodon n
614	20	95.2	125	1	CRCB_AGRT5	Q8ufc8 agrobacteri
615	20	95.2	125	2	Q8TQC6_METAC	Q8tqc6 methanosarc
616	20	95.2	125	2	Q936F9_STAAU	Q936f9 staphylococ
617	20	95.2	125	2	Q5GU44_XANOR	Q5gu44 xanthomonas
618	20	95.2	125	2	Q5G2B7_9HEPC	Q5g2b7 hepatitis c
619	20	95.2	125	2	Q9YLA3_9GEMI	Q9yla3 macroptiliu
620	20	95.2	126	1	URE2_BACPA	P41021 bacillus pa
621	20	95.2	126	2	Q5V566_HALMA	Q5v566 haloarcula
622	20	95.2	126	2	Q4J2G4_AZOVI	Q4j2g4 azotobacter
623	20	95.2	126	2	Q6MPD7_BDEBA	Q6mpd7 bdellovibri
624	20	95.2	127	2	Q8TX75_METKA	Q8tx75 methanopyru
625	20	95.2	127	2	Q56ES6_9CAUD	Q56es6 aeromonas p
626	20	95.2	127	2	Q9L5Q7_SALTI	Q9l5q7 salmonella
627	20	95.2	127	2	Q96622_9GEMI	Q96622 african tom
628	20	95.2	128	2	Q9Y9Q1_AERPE	Q9y9q1 aeropyrum p
629	20	95.2	128	2	Q5AX09_EMENI	Q5ax09 aspergillus
630	20	95.2	128	2	Q6S8S1_PLAFA	Q6s8s1 plasmodium
631	20	95.2	128	2	Q4X8A8_PLACH	Q4x8a8 plasmodium
632	20	95.2	128	2	Q4XHD1_PLACH	Q4xhd1 plasmodium
633	20	95.2	128	2	O04349_ARATH	O04349 arabidopsis
634	20	95.2	128	2	Q87I54_VIBPA	Q87i54 vibrio para
635	20	95.2	128	2	Q4TEC9_TETNG	Q4tec9 tetraodon n
636	20	95.2	129	2	Q8U2C3_PYRFU	Q8u2c3 pyrococcus
637	20	95.2	129	2	Q8U2P6_PYRFU	Q8u2p6 pyrococcus
638	20	95.2	129	2	O58943_PYRHO	O58943 pyrococcus
639	20	95.2	129	2	Q4IF60_GIBZE	Q4if60 gibberella
640	20	95.2	129	2	Q26382_LYMST	Q26382 lymnaea sta
641	20	95.2	129	2	Q4ZMJ3_PSESY	Q4zmj3 pseudomonas
642	20	95.2	129	2	Q5P129_AZOSE	Q5p129 azoarcus sp
643	20	95.2	130	1	V29K_TRVTC	P05074 tobacco rat
644	20	95.2	130	2	Q9HKA4_THEAC	Q9hka4 thermoplasm
645	20	95.2	130	2	Q60LI6_CAEBR	Q60li6 caenorhabdi
646	20	95.2	130	2	Q9N5E4_CAEEL	Q9n5e4 caenorhabdi
647	20	95.2	130	2	Q9I2D5_PSEAE	Q9i2d5 pseudomonas
648	20	95.2	130	2	Q62J42_BURMA	Q62j42 burkholderi
649	20	95.2	130	2	P87527_9RETR	P87527 bovine immu
650	20	95.2	130	2	P87529_9RETR	P87529 bovine immu
651	20	95.2	130	2	P90289_9RETR	P90289 bovine immu
652	20	95.2	131	2	Q64DL3_9ARCH	Q64dl3 uncultured
653	20	95.2	131	2	O62532_DROME	O62532 drosophila
654	20	95.2	131	2	Q04297_9DIPT	Q04297 scaptomyza
655	20	95.2	131	2	Q5YA65_9CAUD	Q5ya65 bacillus cl
656	20	95.2	131	2	Q7Y0V7_ORYSA	Q7y0v7 oryza sativ
657	20	95.2	131	2	Q6T1G0_RHIME	Q6t1g0 rhizobium m
658	20	95.2	131	2	Q57TJ3_SALCH	Q57tj3 salmonella
659	20	95.2	131	2	Q5PIL6_SALPA	Q5pil6 salmonella
660	20	95.2	131	2	Q7CR90_SALTY	Q7cr90 salmonella
661	20	95.2	131	2	Q8XFZ2_SALTI	Q8xfz2 salmonella
662	20	95.2	131	2	Q8JVS0_9POTV	Q8jvs0 soybean mos
663	20	95.2	131	2	P90358_9RETR	P90358 bovine immu
664	20	95.2	132	2	Q51RA4_MAGGR	Q51ra4 magnaporthe
665	20	95.2	132	2	Q6TM68_BPD31	Q6tm68 bacteriopha
666	20	95.2	132	2	Q857I2_BPMB2	Q857i2 mycobacteri
667	20	95.2	132	2	Q6QET8_MYCCE	Q6qet8 mycobacteri
668	20	95.2	132	2	Q6QEU0_9MYCO	Q6qeu0 mycobacteri
669	20	95.2	132	2	Q5HWR5_CAMJR	Q5hwr5 campylobact
670	20	95.2	132	2	Q8EKD1_SHEON	Q8ekd1 shewanella
671	20	95.2	133	2	Q8PZV3_METMA	Q8pzv3 methanosarc
672	20	95.2	133	2	Q8TLA9_METAC	Q8tla9 methanosarc
673	20	95.2	133	2	Q4UCV2_THEAN	Q4ucv2 theileria a
674	20	95.2	133	2	Q9BK39_DRONE	Q9bk39 drosophila
675	20	95.2	133	2	Q81AM8_BACCR	Q81am8 bacillus ce
676	20	95.2	133	2	Q88VV5_LACPL	Q88vv5 lactobacill
677	20	95.2	134	2	Q45786_BACTU	Q45786 bacillus th
678	20	95.2	134	2	Q49007_MYCCA	Q49007 mycoplasma
679	20	95.2	134	2	Q8FAZ2_ECOL6	Q8faz2 escherichia
680	20	95.2	134	2	Q92N40_RHIME	Q92n40 rhizobium m
681	20	95.2	134	2	Q64G53_9HIV1	Q64g53 human immun
682	20	95.2	135	2	Q86XT0_HUMAN	Q86xt0 homo sapien
683	20	95.2	135	2	Q9UDC2_HUMAN	Q9udc2 homo sapien
684	20	95.2	135	2	Q5DFZ2_SCHJA	Q5dfz2 schistosoma
685	20	95.2	135	2	O17529_CAEEL	O17529 caenorhabdi
686	20	95.2	135	2	Q949K4_LYCES	Q949k4 lycopersico
687	20	95.2	135	2	P74277_SYNY3	P74277 synechocyst
688	20	95.2	135	2	Q8D3E9_WIGBR	Q8d3e9 wiggleswort

689	20	95.2	135	2	Q8R3H3_MOUSE	Q8r3h3	mus musculu
690	20	95.2	136	1	RL17_RICPR	Q9zct0	rickettsia
691	20	95.2	136	1	S25BP_HUMAN	O95295	homo sapien
692	20	95.2	136	1	S25BP_MOUSE	Q9z266	mus musculu
693	20	95.2	136	1	S25BP_RAT	P60192	rattus norv
694	20	95.2	136	2	Q8PSY3_METMA	Q8psy3	methanosarc
695	20	95.2	136	2	Q5SXU8_HUMAN	Q5sxu8	homo sapien
696	20	95.2	136	2	Q4HHI1_CAMCO	Q4hhi1	campylobact
697	20	95.2	136	2	Q62DY3_BURMA	Q62dy3	burkholderi
698	20	95.2	137	2	Q6CTN5_KLULA	Q6ctn5	kluyveromyc
699	20	95.2	137	2	Q7YRK4_TARSY	Q7yrk4	tarsius syr
700	20	95.2	137	2	Q9S127_ECOLI	Q9s127	escherichia
701	20	95.2	137	2	Q8GAJ5_ARTNI	Q8gaj5	arthrobacte
702	20	95.2	137	2	Q72M81_LEPIC	Q72m81	leptospira
703	20	95.2	137	2	Q8EYR4_LEPIN	Q8eyr4	leptospira
704	20	95.2	137	2	Q68WA3_RICTY	Q68wa3	rickettsia
705	20	95.2	138	2	Q50ZP0_ENTHI	Q50zp0	entamoeba h
706	20	95.2	138	2	Q88HL5_PSEPK	Q88hl5	pseudomonas
707	20	95.2	138	2	Q9JUH9_NEIMA	Q9juh9	neisseria m
708	20	95.2	138	2	Q6T372_9PARA	Q6t372	newcastle d
709	20	95.2	138	2	Q4RVV8_TETNG	Q4rvv8	tetraodon n
710	20	95.2	139	1	RBS_OLILU	P14961	olisthodisc
711	20	95.2	139	2	Q8TP87_METAC	Q8tp87	methanosarc
712	20	95.2	139	2	Q97Y82_SULSO	Q97y82	sulfolobus
713	20	95.2	139	2	Q4YH82_PLABE	Q4yh82	plasmodium
714	20	95.2	139	2	Q75NF9_9STRA	Q75nf9	heterosigma
715	20	95.2	139	2	Q6HVQ3_BACAN	Q6hvq3	bacillus an
716	20	95.2	139	2	Q4MHW2_BACCE	Q4mhw2	bacillus ce
717	20	95.2	139	2	Q6HFR2_BACHK	Q6hfr2	bacillus th
718	20	95.2	139	2	Q733V9_BACC1	Q733v9	bacillus ce
719	20	95.2	139	2	Q9KZZ9_STRCO	Q9kzz9	streptomyce
720	20	95.2	139	2	Q637T7_BACCZ	Q637t7	bacillus ce
721	20	95.2	139	2	Q5L4H9_TYLCV	Q514h9	tomato yell
722	20	95.2	139	2	Q5L4I1_TYLCV	Q514i1	tomato yell
723	20	95.2	139	2	Q5L4I3_TYLCV	Q514i3	tomato yell
724	20	95.2	139	2	Q6J4T6_9FLAV	Q6j4t6	tick-borne
725	20	95.2	140	2	Q8TSR6_METAC	Q8tsr6	methanosarc
726	20	95.2	140	2	Q8ZYL8_PYRAE	Q8zyl8	pyrobaculum
727	20	95.2	140	2	Q8T7A5_AUTIO	Q8t7a5	autoneris i
728	20	95.2	140	2	Q9MCK1_9CAUD	Q9mck1	streptococc
729	20	95.2	140	2	Q9XJA4_9CAUD	Q9xja4	streptococc
730	20	95.2	140	2	Q8S3W2_CUCSA	Q8s3w2	cucumis sat
731	20	95.2	140	2	Q5LED0_BACFN	Q5led0	bacteroides
732	20	95.2	140	2	Q64VI3_BACFR	Q64vi3	bacteroides
733	20	95.2	140	2	Q88ML7_PSEPK	Q88ml7	pseudomonas
734	20	95.2	140	2	Q5LW31_SILPO	Q5lw31	silicibacte
735	20	95.2	140	2	Q8X8S0_ECO57	Q8x8s0	escherichia
736	20	95.2	140	2	Q83LA5_SHIFL	Q831a5	shigella fl
737	20	95.2	140	2	Q83LQ6_SHIFL	Q831q6	shigella fl
738	20	95.2	141	1	RUVX_PSEPK	Q88d32	pseudomonas
739	20	95.2	141	2	Q8T7A0_9NEOP	Q8t7a0	eacles impe
740	20	95.2	141	2	Q9BNZ3_DRYRU	Q9bnz3	dryocampa r
741	20	95.2	141	2	O64278_9CAUD	O64278	streptococc
742	20	95.2	141	2	Q4K9C8_PSEF5	Q4k9c8	pseudomonas
743	20	95.2	141	2	Q6LRD8_PHOPR	Q6lrd8	photobacter
744	20	95.2	141	2	Q71SZ2_9SCOM	Q71sz2	scombridae
745	20	95.2	141	2	Q71SZ3_CYPCA	Q71sz3	cyprinus ca
746	20	95.2	141	2	Q4RS92_TETNG	Q4rs92	tetraodon n
747	20	95.2	142	2	Q50TF4_ENTHI	Q50tf4	entamoeba h
748	20	95.2	142	2	Q5YEW2_FRAAN	Q5yew2	fragaria an
749	20	95.2	142	2	Q82FS2_STRAW	Q82fs2	streptomyce
750	20	95.2	142	2	O90153_9FLAV	O90153	hepatitis g
751	20	95.2	143	2	Q6MYB1_ASPFU	Q6myb1	aspergillus
752	20	95.2	143	2	Q584P6_9TRYP	Q584p6	trypanosoma
753	20	95.2	143	2	Q851L7_ORYSA	Q851l7	oryza sativ
754	20	95.2	143	2	Q9FW54_ORYSA	Q9fw54	oryza sativ
755	20	95.2	143	2	Q53JK5_ORYSA	Q53jk5	oryza sativ
756	20	95.2	143	2	Q8GBJ6_9BACT	Q8gbj6	uncultured
757	20	95.2	143	2	Q5H445_XANOR	Q5h445	xanthomonas
758	20	95.2	143	2	Q9AK79_STRCO	Q9ak79	streptomyce
759	20	95.2	144	1	RPI_ECOLI	P13809	escherichia
760	20	95.2	144	1	RS15_CHICK	P62846	gallus gall
761	20	95.2	144	1	RS15_HUMAN	P62841	homo sapien
762	20	95.2	144	1	RS15_MESAU	P62842	mesocricetu
763	20	95.2	144	1	RS15_MOUSE	P62843	mus musculu
764	20	95.2	144	1	RS15_PIG	P62844	sus scrofa
765	20	95.2	144	1	RS15_RAT	P62845	rattus norv
766	20	95.2	144	1	RS15_XENLA	P20342	xenopus lae
767	20	95.2	144	1	RUVX_PSEAE	Q9i699	pseudomonas
768	20	95.2	144	2	Q938K5_9CAUD	Q938k5	temperate p
769	20	95.2	144	2	Q655Q4_ORYSA	Q655q4	oryza sativ
770	20	95.2	144	2	Q79XT6_STRP3	Q79xt6	streptococc
771	20	95.2	144	2	Q5XAS5_STRP6	Q5xas5	streptococc
772	20	95.2	144	2	Q9A0N2_STRPY	Q9a0n2	streptococc
773	20	95.2	144	2	Q8NZS2_STRP8	Q8nzs2	streptococc
774	20	95.2	144	2	Q7UT05_RHOBA	Q7ut05	rhodopirell
775	20	95.2	145	1	COX5A_MOUSE	P12787	mus musculu
776	20	95.2	145	1	RS15_XIPMA	P70066	xiphophorus
777	20	95.2	145	1	RUVX_PSEFL	Q9f4i8	pseudomonas
778	20	95.2	145	2	Q5BZ98_SCHJA	Q5bz98	schistosoma
779	20	95.2	145	2	Q56K10_BOVIN	Q56k10	bos taurus
780	20	95.2	145	2	Q5RDI7_PONPY	Q5rdi7	pongo pygma
781	20	95.2	145	2	Q4K4E4_PSEF5	Q4k4e4	pseudomonas
782	20	95.2	145	2	Q7UZN2_PROMP	Q7uzn2	prochloroco
783	20	95.2	145	2	Q6D6I2_ERWCT	Q6d6i2	erwinia car
784	20	95.2	145	2	Q7NPJ0_GLOVI	Q7npj0	gloeobacter
785	20	95.2	145	2	Q8FTB4_COREF	Q8ftb4	corynebacte
786	20	95.2	145	2	Q63JB7_BURPS	Q63jb7	burkholderi
787	20	95.2	145	2	Q52KC9_MOUSE	Q52kc9	mus musculu
788	20	95.2	145	2	Q5RJJ6_XENTR	Q5rjj6	xenopus tro
789	20	95.2	145	2	Q642P0_XENLA	Q642p0	xenopus lae
790	20	95.2	145	2	Q6DGW6_BRARE	Q6dgw6	brachydanio
791	20	95.2	145	2	Q6Q421_BRARE	Q6q421	brachydanio
792	20	95.2	145	2	Q71U13_LAPHA	Q71u13	lapemis har
793	20	95.2	145	2	Q90YQ9_ICTPU	Q90yq9	ictalurus p
794	20	95.2	145	2	Q9YGQ0_SALSA	Q9ygq0	salmo salar
795	20	95.2	146	1	COX5A_RAT	P11240	rattus norv
796	20	95.2	146	2	Q8TJL1_METAC	Q8tjl1	methanosarc
797	20	95.2	146	2	Q5V2B3_HALMA	Q5v2b3	haloarcula
798	20	95.2	146	2	Q8ZSS3_PYRAE	Q8zss3	pyrobaculum
799	20	95.2	146	2	Q9SHZ9_ARATH	Q9shz9	arabidopsis
800	20	95.2	146	2	Q89YB2_BRAJA	Q89yb2	bradyrhizob
801	20	95.2	146	2	O33703_STRPN	O33703	streptococc
802	20	95.2	146	2	Q9D2W1_MOUSE	Q9d2w1	mus musculu
803	20	95.2	147	2	Q7RUW8_NEUCR	Q7ruw8	neurospora
804	20	95.2	147	2	Q5CEJ1_CRYHO	Q5cej1	cryptospori
805	20	95.2	147	2	Q6I585_ORYSA	Q6i585	oryza sativ
806	20	95.2	147	2	Q8DNF8_STRR6	Q8dnf8	streptococc
807	20	95.2	147	2	Q97NS3_STRPN	Q97ns3	streptococc
808	20	95.2	147	2	Q6N6L2_RHOPA	Q6n612	rhodopseudo
809	20	95.2	147	2	Q7URA3_RHOBA	Q7ura3	rhodopirell
810	20	95.2	148	1	YHT6_YEAST	P38839	saccharomyc
811	20	95.2	148	2	P87290_YEAST	P87290	saccharomyc
812	20	95.2	148	2	Q706P2_PSEPU	Q706p2	pseudomonas
813	20	95.2	148	2	Q5H6P1_XANOR	Q5h6p1	xanthomonas
814	20	95.2	148	2	Q4UV34_XANCP	Q4uv34	xanthomonas
815	20	95.2	148	2	Q6IP26_XENLA	Q6ip26	xenopus lae
816	20	95.2	148	2	P87528_9RETR	P87528	bovine immu
817	20	95.2	148	2	P87530_9RETR	P87530	bovine immu
818	20	95.2	149	1	Y1039_BORPE	Q7vz88	bordetella
819	20	95.2	149	2	Q5C1G8_SCHJA	Q5c1g8	schistosoma
820	20	95.2	149	2	Q5QC93_AEDAE	Q5qc93	aedes aegyp
821	20	95.2	149	2	Q5MIR4_AEDAL	Q5mir4	aedes albop
822	20	95.2	149	2	Q6S7F8_9BACT	Q6s7f8	uncultured
823	20	95.2	149	2	Q6S7P1_9BACT	Q6s7p1	uncultured
824	20	95.2	149	2	Q6S7P5_9BACT	Q6s7p5	uncultured
825	20	95.2	149	2	Q6S7P8_9BACT	Q6s7p8	uncultured
826	20	95.2	149	2	Q6S7P9_9BACT	Q6s7p9	uncultured
827	20	95.2	149	2	Q6S7Q2_9BACT	Q6s7q2	uncultured
828	20	95.2	149	2	Q6S7Q9_9BACT	Q6s7q9	uncultured
829	20	95.2	149	2	Q8RAH8_THETN	Q8rah8	thermoanaer
830	20	95.2	149	2	Q97DW3_CLOAB	Q97dw3	clostridium
831	20	95.2	149	2	Q4S8B4_TETNG	Q4s8b4	tetraodon n
832	20	95.2	150	1	COX5A_HUMAN	P20674	homo sapien
833	20	95.2	150	2	Q6R656_MELLI	Q6r656	melampsora
834	20	95.2	150	2	Q8TB65_HUMAN	Q8tb65	homo sapien

835	20	95.2	150	2	Q5DBT0_SCHJA	Q5dbt0	schistosoma
836	20	95.2	150	2	Q4Z529_PLABE	Q4z529	plasmodium
837	20	95.2	150	2	Q53CF8_MACMU	Q53cf8	macaca mula
838	20	95.2	150	2	Q53CG1_SAISC	Q53cg1	saimiri sci
839	20	95.2	150	2	Q5HT49_CAMJR	Q5ht49	campylobact
840	20	95.2	150	2	Q9PMT8_CAMJE	Q9pmt8	campylobact
841	20	95.2	150	2	Q884K7_PSESM	Q884k7	pseudomonas
842	20	95.2	150	2	Q67NK6_SYMTH	Q67nk6	symbiobacte
843	20	95.2	150	2	O09776_9CALI	O09776	reptile cal
844	20	95.2	150	2	Q8Q4N0_SIVCZ	Q8q4n0	chimpanzee
845	20	95.2	151	1	RUVX_NEIG1	Q5f936	neisseria g
846	20	95.2	151	1	RUVX_NEIMA	Q9ju11	neisseria m
847	20	95.2	151	1	RUVX_NEIMB	Q9jz16	neisseria m
848	20	95.2	151	1	Y903_AQUAE	O67050	aquifex aeo
849	20	95.2	151	2	Q4WS67_ASPFU	Q4ws67	aspergillus
850	20	95.2	151	2	Q5S1X4_IXOSC	Q5s1x4	ixodes scap
851	20	95.2	151	2	Q25015_9VEST	Q25015	haliotis au
852	20	95.2	151	2	Q9KKH6_YEREN	Q9kkh6	yersinia en
853	20	95.2	151	2	Q6XA36_9PARA	Q6xa36	newcastle d
854	20	95.2	151	2	Q6XA38_9PARA	Q6xa38	newcastle d
855	20	95.2	151	2	Q6XA40_9PARA	Q6xa40	newcastle d
856	20	95.2	151	2	Q6XA42_9PARA	Q6xa42	newcastle d
857	20	95.2	151	2	Q6XA44_9PARA	Q6xa44	newcastle d
858	20	95.2	151	2	Q6XA46_9PARA	Q6xa46	newcastle d
859	20	95.2	151	2	Q6XA48_9PARA	Q6xa48	newcastle d
860	20	95.2	151	2	Q6XA50_9PARA	Q6xa50	newcastle d
861	20	95.2	151	2	Q6XA52_9PARA	Q6xa52	newcastle d
862	20	95.2	151	2	Q6XA54_9PARA	Q6xa54	newcastle d
863	20	95.2	151	2	Q6XA56_9PARA	Q6xa56	newcastle d
864	20	95.2	151	2	Q6XA58_9PARA	Q6xa58	newcastle d
865	20	95.2	151	2	Q6XA60_9PARA	Q6xa60	newcastle d
866	20	95.2	151	2	Q6XA62_9PARA	Q6xa62	newcastle d
867	20	95.2	151	2	Q6NYP1_BRARE	Q6nyp1	brachydanio
868	20	95.2	151	2	Q6PBX0_BRARE	Q6pbx0	brachydanio
869	20	95.2	151	2	Q4T9H6_TETNG	Q4t9h6	tetraodon n
870	20	95.2	151	2	Q4SMA7_TETNG	Q4sma7	tetraodon n
871	20	95.2	152	1	COX5A_BOVIN	P00426	bos taurus
872	20	95.2	152	2	Q4IH16_GIBZE	Q4ih16	gibberella
873	20	95.2	152	2	Q7PSR9_ANOGA	Q7psr9	anopheles g
874	20	95.2	152	2	Q6IJ02_DROME	Q6ij02	drosophila
875	20	95.2	152	2	Q50M98_ENTHI	Q50m98	entamoeba h
876	20	95.2	152	2	Q5XD03_STRP6	Q5xd03	streptococc
877	20	95.2	152	2	Q9A0K8_STRPY	Q9a0k8	streptococc
878	20	95.2	152	2	Q745S5_MYCPA	Q745s5	mycobacteri
879	20	95.2	152	2	Q7TVS1_MYCBO	Q7tvs1	mycobacteri
880	20	95.2	152	2	Q8P1M2_STRP8	Q8p1m2	streptococc
881	20	95.2	152	2	O69719_MYCTU	O69719	mycobacteri
882	20	95.2	152	2	Q7CFA9_STRP3	Q7cfa9	streptococc
883	20	95.2	153	2	Q6B867_9ACAR	Q6b867	ixodes paci
884	20	95.2	153	2	Q9Y1I7_9VEST	Q9y1i7	tegula aure
885	20	95.2	153	2	Q6ILV3_DROME	Q6ilv3	drosophila
886	20	95.2	153	2	Q50VS3_ENTHI	Q50vs3	entamoeba h
887	20	95.2	153	2	O64289_9CAUD	O64289	streptococc
888	20	95.2	153	2	Q5FTK5_GLUOX	Q5ftk5	gluconobact
889	20	95.2	153	2	Q7U816_SYNPX	Q7u816	synechococc
890	20	95.2	154	1	Y726_AZOSE	Q5p762	azoarcus sp
891	20	95.2	154	2	Q61VP3_CAEBR	Q61vp3	caenorhabdi
892	20	95.2	154	2	Q5C1K9_SCHJA	Q5c1k9	schistosoma
893	20	95.2	154	2	Q8GMN3_AERSA	Q8gmn3	aeromonas s
894	20	95.2	154	2	Q5P0G9_AZOSE	Q5p0g9	azoarcus sp
895	20	95.2	154	2	Q8EMU1_OCEIH	Q8emu1	oceanobacil
896	20	95.2	154	2	Q6T343_9PARA	Q6t343	newcastle d
897	20	95.2	154	2	Q6T370_9PARA	Q6t370	newcastle d
898	20	95.2	155	2	Q4XEJ2_PLACH	Q4xej2	plasmodium
899	20	95.2	155	2	Q7VS68_BORPE	Q7vs68	bordetella
900	20	95.2	155	2	Q8XPQ0_RALSO	Q8xpq0	ralstonia s
901	20	95.2	156	1	ELYS_HALCR	Q01380	haliotis cr
902	20	95.2	156	2	Q55W00_CRYNE	Q55w00	cryptococcu
903	20	95.2	156	2	Q5KK91_CRYNE	Q5kk91	cryptococcu
904	20	95.2	156	2	Q7RTF1_PLAYO	Q7rtf1	plasmodium
905	20	95.2	156	2	Q7P329_FUSNV	Q7p329	fusobacteri
906	20	95.2	156	2	Q5GV41_XANOR	Q5gv41	xanthomonas
907	20	95.2	156	2	Q5GZ78_XANOR	Q5gz78	xanthomonas
908	20	95.2	156	2	Q4J347_AZOVI	Q4j347	azotobacter
909	20	95.2	156	2	Q87A89_XYLFT	Q87a89	xylella fas
910	20	95.2	156	2	Q8A556_BACTN	Q8a556	bacteroides
911	20	95.2	156	2	Q88A24_PSESM	Q88a24	pseudomonas
912	20	95.2	157	1	Y1139_BORPA	Q7wb68	bordetella
913	20	95.2	157	1	Y1355_BORBR	Q7wmn5	bordetella
914	20	95.2	157	2	Q9VQL3_DROME	Q9vql3	drosophila
915	20	95.2	157	2	Q6ZDG7_ORYSA	Q6zdg7	oryza sativ
916	20	95.2	157	2	Q6ZKY5_ORYSA	Q6zky5	oryza sativ
917	20	95.2	157	2	Q5F7S5_NEIG1	Q5f7s5	neisseria g
918	20	95.2	157	2	Q7NPF8_GLOVI	Q7npf8	gloeobacter
919	20	95.2	157	2	Q7V7E2_PROMM	Q7v7e2	prochloroco
920	20	95.2	158	2	Q55CP5_DICDI	Q55cp5	dictyosteli
921	20	95.2	158	2	Q5TWC0_ANOGA	Q5twc0	anopheles g
922	20	95.2	158	2	Q6YDF3_9FABA	Q6ydf3	arachis car
923	20	95.2	158	2	Q6YWY6_ORYSA	Q6ywy6	oryza sativ
924	20	95.2	158	2	Q5GUC0_XANOR	Q5guc0	xanthomonas
925	20	95.2	158	2	Q5GVP4_XANOR	Q5gvp4	xanthomonas
926	20	95.2	158	2	Q5H030_XANOR	Q5h030	xanthomonas
927	20	95.2	158	2	Q5H119_XANOR	Q5h119	xanthomonas
928	20	95.2	158	2	Q5H3K2_XANOR	Q5h3k2	xanthomonas
929	20	95.2	158	2	Q5H3S0_XANOR	Q5h3s0	xanthomonas
930	20	95.2	158	2	Q5H648_XANOR	Q5h648	xanthomonas
931	20	95.2	158	2	Q6MPD6_BDEBA	Q6mpd6	bdellovibri
932	20	95.2	159	2	Q6BZQ8_YARLI	Q6bzq8	yarrowia li
933	20	95.2	159	2	Q5BXJ9_SCHJA	Q5bxj9	schistosoma
934	20	95.2	159	2	Q4YH56_PLABE	Q4yh56	plasmodium
935	20	95.2	159	2	Q6VTB3_9BACT	Q6vtb3	symbiont ba
936	20	95.2	159	2	Q4LBL9_PSESH	Q4lbl9	pseudomonas
937	20	95.2	159	2	Q4KD78_PSEF5	Q4kd78	pseudomonas
938	20	95.2	159	2	Q5N419_SYNP6	Q5n419	synechococc
939	20	95.2	159	2	Q5Z3B1_NOCFA	Q5z3b1	nocardia fa
940	20	95.2	159	2	Q8PHI9_XANAC	Q8phi9	xanthomonas
941	20	95.2	159	2	Q6XA26_9PARA	Q6xa26	newcastle d
942	20	95.2	159	2	Q6XA28_9PARA	Q6xa28	newcastle d
943	20	95.2	159	2	Q6XA30_9PARA	Q6xa30	newcastle d
944	20	95.2	159	2	Q6XA34_9PARA	Q6xa34	newcastle d
945	20	95.2	160	1	PHAA_AGLNE	P28555	aglaothamni
946	20	95.2	160	1	Y650_TREPA	O83656	treponema p
947	20	95.2	160	2	Q8PXB2_METMA	Q8pxb2	methanosarc
948	20	95.2	160	2	Q75EX7_ASHGO	Q75ex7	ashbya goss
949	20	95.2	160	2	Q7PN92_ANOGA	Q7pn92	anopheles g
950	20	95.2	160	2	Q8W2S9_ORYSA	Q8w2s9	oryza sativ
951	20	95.2	160	2	Q5H107_XANOR	Q5h107	xanthomonas
952	20	95.2	160	2	Q6F6S8_ACIAD	Q6f6s8	acinetobact
953	20	95.2	160	2	Q7U4N4_SYNPX	Q7u4n4	synechococc
954	20	95.2	160	2	Q6TNQ4_BRARE	Q6tnq4	brachydanio
955	20	95.2	161	1	MLR_DICDI	P13833	dictyosteli
956	20	95.2	161	2	Q552B5_DICDI	Q552b5	dictyosteli
957	20	95.2	161	2	Q5Q9W9_NICGL	Q5q9w9	nicotiana g
958	20	95.2	161	2	Q4KB46_PSEF5	Q4kb46	pseudomonas
959	20	95.2	161	2	Q9ZKH9_HELPJ	Q9zkh9	helicobacte
960	20	95.2	162	1	CAV2_CANFA	O46550	canis famil
961	20	95.2	162	2	Q51A31_ENTHI	Q51a31	entamoeba h
962	20	95.2	162	2	Q4Q087_LEIMA	Q4q087	leishmania
963	20	95.2	162	2	Q4FD05_PIG	Q4fd05	sus scrofa
964	20	95.2	162	2	O69752_PSEAE	O69752	pseudomonas
965	20	95.2	162	2	Q9HWH1_PSEAE	Q9hwh1	pseudomonas
966	20	95.2	162	2	Q5NYT4_AZOSE	Q5nyt4	azoarcus sp
967	20	95.2	162	2	Q62582_PERLE	Q62582	peromyscus
968	20	95.2	162	2	Q6ZWQ5_MOUSE	Q6zwq5	mus musculu
969	20	95.2	162	2	Q5ZHT0_CHICK	Q5zht0	gallus gall
970	20	95.2	162	2	Q7ZY55_XENLA	Q7zy55	xenopus lae
971	20	95.2	162	2	Q803J1_BRARE	Q803j1	brachydanio
972	20	95.2	162	2	Q5EBE4_XENTR	Q5ebe4	xenopus tro
973	20	95.2	162	2	Q6GPZ0_XENLA	Q6gpz0	xenopus lae
974	20	95.2	162	2	Q4RAF8_TETNG	Q4raf8	tetraodon n
975	20	95.2	163	2	Q5DHH7_SCHJA	Q5dhh7	schistosoma
976	20	95.2	163	2	Q9W372_DROME	Q9w372	drosophila
977	20	95.2	163	2	Q4XQC8_PLACH	Q4xqc8	plasmodium
978	20	95.2	163	2	Q9XUF0_CAEEL	Q9xuf0	caenorhabdi
979	20	95.2	163	2	Q9FPV3_PETIN	Q9fpv3	petunia int
980	20	95.2	163	2	Q8DG89_SYNEL	Q8dg89	synechococc

981	20	95.2	164	2	Q8N112_HUMAN	Q8n112 homo sapien
982	20	95.2	164	2	Q6EPZ1_ORYSA	Q6epz1 oryza sativ
983	20	95.2	164	2	Q7XHL4_ORYSA	Q7xhl4 oryza sativ
984	20	95.2	164	2	Q8XSK9_RALSO	Q8xsk9 ralstonia s
985	20	95.2	164	2	Q9KKX4_VIBCH	Q9kkx4 vibrio chol
986	20	95.2	164	2	Q67RM1_SYMTH	Q67rm1 symbiobacte
987	20	95.2	164	2	Q5F228_MOUSE	Q5f228 mus musculu
988	20	95.2	164	2	Q5RH22_BRARE	Q5rh22 brachydanio
989	20	95.2	165	1	SNX12_MOUSE	O70493 mus musculu
990	20	95.2	165	2	Q513T7_ENTHI	Q513t7 entamoeba h
991	20	95.2	165	2	Q4YIV0_PLABE	Q4yiv0 plasmodium
992	20	95.2	165	2	Q71IT5_LACDL	Q71it5 lactobacill
993	20	95.2	165	2	Q4J2V2_AZOVI	Q4j2v2 azotobacter
994	20	95.2	165	2	Q8D7T9_VIBVU	Q8d7t9 vibrio vuln
995	20	95.2	165	2	Q928B3_LISIN	Q928b3 listeria in
996	20	95.2	165	2	Q986B3_RHILO	Q986b3 rhizobium l
997	20	95.2	166	2	Q9BK38_9DIPT	Q9bk38 lordiphosa
998	20	95.2	166	2	Q56C19_9CAUD	Q56c19 enterobacte
999	20	95.2	166	2	Q8GQ00_PSEAE	Q8gq00 pseudomonas
1000	20	95.2	166	2	Q8DN14_STRR6	Q8dn14 streptococc

ALIGNMENTS

```
RESULT 1
Q23912_DICDI
ID   Q23912_DICDI PRELIMINARY;      PRT;    16 AA.
AC   Q23912;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Nucleoside diphosphate kinase Gip17 (Fragment).
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=90277608; PubMed=2161830;
RA   Lacombe M.-L., Wallet V., Troll H., Veron M.;
RT   "Functional cloning of a nucleoside diphosphate kinase from
RT   Dictyostelium discoideum.";
RL   J. Biol. Chem. 265:10012-10018(1990).
DR   EMBL; M36679; AAA33232.1; -; mRNA.
DR   DictyBase; DDB0185051; ndkC.
DR   GO; GO:0016301; F:kinase activity; IEA.
KW   Kinase.
FT   NON_TER      16     16
SQ   SEQUENCE   16 AA;  1897 MW;  38BE6B475C598B5E CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 16;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          6 RRLN 9


RESULT 2
Q5BX77_SCHJA
ID   Q5BX77_SCHJA PRELIMINARY;      PRT;    27 AA.
AC   Q5BX77;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
OS   Schistosoma japonicum (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
OX   NCBI_TaxID=6182;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Han Z.;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY812109; AAX27998.1; -; mRNA.
DR   InterPro; IPR011989; ARM-like.
KW   Hypothetical protein.
SQ   SEQUENCE   27 AA;  3410 MW;  7E0683D08C942315 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 27;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          6 RRLN 9


RESULT 3
Q50L67_DROBP
ID   Q50L67_DROBP PRELIMINARY;      PRT;    33 AA.
AC   Q50L67;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein CG11779 (Fragment).
GN   Name=CG11779;
OS   Drosophila bipectinata (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=42026;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Wild type B133, and Wild type CJB162;
RA   Nozawa M., Aotsuka T., Tamura K.;
RT   "A novel chimeric gene, sirene, found in the Drosophila bipectinata
RT   species complex: potential of retoroposition with regulatory
RT   sequence.";
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB194439; BAD98229.1; -; Genomic_DNA.
DR   EMBL; AB194442; BAD98232.1; -; Genomic_DNA.
DR   EMBL; AB194429; BAD98204.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER      33     33
SQ   SEQUENCE   33 AA;  4154 MW;  B9FF28FCB8AB796C CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 33;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         25 RRLN 28


RESULT 4
Q50L79_9DIPT
ID   Q50L79_9DIPT PRELIMINARY;      PRT;    33 AA.
AC   Q50L79;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein CG11779 (Fragment).
GN   Name=CG11779;
OS   Drosophila pseudoananassae pseudoananassae.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=65965;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Nozawa M., Aotsuka T., Tamura K.;
RT   "A novel chimeric gene, sirene, found in the Drosophila bipectinata
RT   species complex: potential of retoroposition with regulatory
```

```
RT   sequence.";
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB194433; BAD98216.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER      33     33
SQ   SEQUENCE   33 AA;  4128 MW;  E9FF28FCB8BB686D CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 33;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         25 RRLN 28


RESULT 5
Q50L83_9DIPT
ID   Q50L83_9DIPT PRELIMINARY;      PRT;    33 AA.
AC   Q50L83;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein CG11779 (Fragment).
GN   Name=CG11779;
OS   Drosophila pseudoananassae nigrens.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=296643;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Nozawa M., Aotsuka T., Tamura K.;
RT   "A novel chimeric gene, sirene, found in the Drosophila bipectinata
RT   species complex: potential of retoroposition with regulatory
RT   sequence.";
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB194432; BAD98213.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER      33     33
SQ   SEQUENCE   33 AA;  4128 MW;  E9FF28FCB8BB686D CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 33;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         25 RRLN 28


RESULT 6
Q50L86_9DIPT
ID   Q50L86_9DIPT PRELIMINARY;      PRT;    33 AA.
AC   Q50L86;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein CG11779 (Fragment).
GN   Name=CG11779;
OS   Drosophila malerkotliana pallens.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=296644;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Nozawa M., Aotsuka T., Tamura K.;
RT   "A novel chimeric gene, sirene, found in the Drosophila bipectinata
RT   species complex: potential of retoroposition with regulatory
RT   sequence.";
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB194431; BAD98210.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER      33     33
SQ   SEQUENCE   33 AA;  4154 MW;  E9FF28FCB8AB796C CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 33;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         25 RRLN 28


RESULT 7
Q50L89_9DIPT
ID   Q50L89_9DIPT PRELIMINARY;      PRT;    33 AA.
AC   Q50L89;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein CG11779 (Fragment).
GN   Name=CG11779;
OS   Drosophila malerkotliana malerkotliana.
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=195057;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Nozawa M., Aotsuka T., Tamura K.;
RT   "A novel chimeric gene, sirene, found in the Drosophila bipectinata
RT   species complex: potential of retoroposition with regulatory
RT   sequence.";
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB194430; BAD98207.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER      33     33
SQ   SEQUENCE   33 AA;  4154 MW;  E9FF28FCB8AB796C CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 33;
  Best Local Similarity   100.0%;  Pred. No. 3.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         25 RRLN 28


RESULT 8
Q50L77_DROAN
ID   Q50L77_DROAN PRELIMINARY;      PRT;    34 AA.
AC   Q50L77;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein CG11779 (Fragment).
GN   Name=CG11779;
OS   Drosophila ananassae (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC   Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7217;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Nozawa M., Aotsuka T., Tamura K.;
RT   "A novel chimeric gene, sirene, found in the Drosophila bipectinata
RT   species complex: potential of retoroposition with regulatory
RT   sequence.";
RL   Submitted (NOV-2004) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB194434; BAD98219.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER      34     34
SQ   SEQUENCE   34 AA;  4262 MW;  90CF599A9912FA05 CRC64;
```

```
  Query Match             95.2%;  Score 20;  DB 2;  Length 34;
  Best Local Similarity   100.0%;  Pred. No. 4.1e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         26 RRLN 29


RESULT 9
Q4ZG64_HUMAN
ID   Q4ZG64_HUMAN PRELIMINARY;      PRT;    35 AA.
AC   Q4ZG64;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein SP100 (Fragment).
GN   Name=SP100;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Abbott A., McLellan M., Haub K.;
RT   "The sequence of Homo sapiens BAC clone RP11-69J7.";
RL   Submitted (SEP-1999) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Waterston R.H.;
RL   Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Waterston R.;
RL   Submitted (OCT-2002) to the EMBL/GenBank/DDBJ databases.
RN   [4]
RP   NUCLEOTIDE SEQUENCE.
RA   Wilson R.K.;
RL   Submitted (APR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AC009949; AAX88870.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER      35     35
SQ   SEQUENCE   35 AA;  3729 MW;  0DD1F466F6016B22 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 35;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         11 RRLN 14


RESULT 10
Q888U8_PSESM
ID   Q888U8_PSESM PRELIMINARY;      PRT;    35 AA.
AC   Q888U8;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=PSPTO0918;
OS   Pseudomonas syringae (pv. tomato).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC   Pseudomonadaceae; Pseudomonas.
OX   NCBI_TaxID=323;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=DC3000;
RX   MEDLINE=22834015; PubMed=12928499; DOI=10.1073/pnas.1731982100;
RA   Buell C.R., Joardar V., Lindeberg M., Selengut J., Paulsen I.T.,
RA   Gwinn M.L., Dodson R.J., DeBoy R.T., Durkin A.S., Kolonay J.F.,
RA   Madupu R., Daugherty S.C., Brinkac L.M., Beanan M.J., Haft D.H.,
RA   Nelson W.C., Davidsen T.M., Zafar N., Zhou L., Liu J., Yuan Q.,
RA   Khouri H.M., Fedorova N.B., Tran B., Russell D., Berry K.J.,
RA   Utterback T.R., Van Aken S.E., Feldblyum T.V., D'Ascenzo M.,
RA   Deng W.-L., Ramos A.R., Alfano J.R., Cartinhour S., Chatterjee A.K.,
RA   Delaney T.P., Lazarowitz S.G., Martin G.B., Schneider D.J., Tang X.,
RA   Bender C.L., White O., Fraser C.M., Collmer A.;
RT   "The complete genome sequence of the Arabidopsis and tomato pathogen
RT   Pseudomonas syringae pv. tomato DC3000.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:10181-10186(2003).
DR   EMBL; AE016853; AAO54452.1; -; Genomic_DNA.
DR   TIGR; PSPTO0918; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   35 AA;  4182 MW;  6F44D83F7F895634 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 35;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          3 RRLN 6


RESULT 11
Q4SVA1_TETNG
ID   Q4SVA1_TETNG PRELIMINARY;      PRT;    35 AA.
AC   Q4SVA1;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome undetermined SCAF13770, whole genome shotgun sequence.
DE   (Fragment).
GN   ORFNames=GSTENG00012082001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01013770; CAF95431.1; -; Genomic_DNA.
FT   NON_TER       1      1
SQ   SEQUENCE   35 AA;  3917 MW;  104DEC79CCD12A70 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 35;
  Best Local Similarity   100.0%;  Pred. No. 4.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 RRLN 4
              ||||
Db         26 RRLN 29


RESULT 12
Q4Z5P5_PLABE
ID   Q4Z5P5_PLABE PRELIMINARY;      PRT;    36 AA.
AC   Q4Z5P5;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein (Fragment).
GN   ORFNames=PB100986.00.0;
OS   Plasmodium berghei.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=5821;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Hall N., Karras M., Raine J.D., Carlton J.M., Kooij T.W.A.,
RA   Berriman M., Florens L., Janssen C.S., Pain A., Christophides G.K.,
RA   James K., Rutherford K., Harris B., Harris D., Churcher C.,
RA   Quail M.A., Ormond D., Doggett J., Trueman H.E., Mendoza J.,
RA   Bidwell S.L., Rajandream M.A., Carucci D.J, Yates J.R., Kafatos F.C.,
RA   Janse C.J., Barrell B., Turner C.M.R., Waters A.P., Sinden R.S.;
RT   "A comprehensive survey of the Plasmodium life cycle by genomic,
RT   transcriptomic, and proteomic analyses.";
RL   Science 307:82-86(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAI01000412; CAH94332.1; -; Genomic_DNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   36 AA;  4144 MW;  A481C901A7F72669 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 36;
  Best Local Similarity   100.0%;  Pred. No. 4.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         20 RRLN 23


RESULT 13
Q7ULM0_RHOBA
ID   Q7ULM0_RHOBA PRELIMINARY;      PRT;    36 AA.
AC   Q7ULM0;
DT   01-OCT-2003 (TrEMBLrel. 25, Created)
DT   01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=RB9415;
OS   Rhodopirellula baltica.
OC   Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC   Planctomycetaceae; Pirellula.
OX   NCBI_TaxID=117;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=1;
RX   MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA   Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA   Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA   Schlesner H., Amann R., Reinhardt R.;
RT   "Complete genome sequence of the marine planctomycete Pirellula sp.
RT   strain 1.";
RL   Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303(2003).
DR   EMBL; BX294149; CAD76249.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   36 AA;  4076 MW;  EEBB6487A9EC415D CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 36;
  Best Local Similarity   100.0%;  Pred. No. 4.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         26 RRLN 29


RESULT 14
Q7SHZ1_NEUCR
ID   Q7SHZ1_NEUCR PRELIMINARY;      PRT;    37 AA.
AC   Q7SHZ1;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Predicted protein.
GN   Name=NCU00661.1;
OS   Neurospora crassa.
OC   Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC   Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX   NCBI_TaxID=5141;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=OR74A;
RA   Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
RA   Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
RA   Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
RA   Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
RA   Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
RA   Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA   Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
RA   Kamal M., Kamvysselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA   Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA   Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
RA   DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA   Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA   Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
RA   Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
RT   "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL   Nature 0:0-0(2003).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABX01000001; EAA36565.1; -; Genomic_DNA.
SQ   SEQUENCE   37 AA;  4165 MW;  D6339C3DB0B98F60 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 4.5e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         16 RRLN 19


RESULT 15
Q8F419_LEPIN
ID   Q8F419_LEPIN PRELIMINARY;      PRT;    37 AA.
AC   Q8F419;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=LA2228;
OS   Leptospira interrogans.
OC   Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX   NCBI_TaxID=173;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RX   MEDLINE=22598143; PubMed=12712204; DOI=10.1038/nature01597;
RA   Ren S.-X., Fu G., Jiang X.-G., Zeng R., Miao Y.-G., Xu H.,
RA   Zhang Y.-X., Xiong H., Lu G., Lu L.-F., Jiang H.-Q., Jia J., Tu Y.-F.,
```

```
RA   Jiang J.-X., Gu W.-Y., Zhang Y.-Q., Cai Z., Sheng H.-H., Yin H.-F.,
RA   Zhang Y., Zhu G.-F., Wan M., Huang H.-L., Qian Z., Wang S.-Y., Ma W.,
RA   Yao Z.-J., Shen Y., Qiang B.-Q., Xia Q.-C., Guo X.-K., Danchin A.,
RA   Saint Girons I., Somerville R.L., Wen Y.-M., Shi M.-H., Chen Z.,
RA   Xu J.-G., Zhao G.-P.;
RT   "Unique physiological and pathogenic features of Leptospira
RT   interrogans revealed by whole-genome sequencing.";
RL   Nature 422:888-893(2003).
DR   EMBL; AE011394; AAN49427.1; -; Genomic_DNA.
KW   Complete proteome.
SQ   SEQUENCE   37 AA;  4152 MW;  CC11BD48D35A495C CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 37;
  Best Local Similarity   100.0%;  Pred. No. 4.5e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          5 RRLN 8


RESULT 16
Q4TTJ0_STRPU
ID   Q4TTJ0_STRPU PRELIMINARY;      PRT;    39 AA.
AC   Q4TTJ0;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Extracellular protein (Fragment).
GN   Name=endo16;
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15937122;
RA   Balhoff J.P., Wray G.A.;
RT   "Evolutionary analysis of the well characterized endo16 promoter
RT   reveals substantial variation within functional sites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 102:8591-8596(2005).
DR   EMBL; DQ066808; AAY51737.1; -; Genomic_DNA.
FT   NON_TER      39     39
SQ   SEQUENCE   39 AA;  4429 MW;  2AA9F34D78AADA6F CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 17
Q4TTJ2_STRPU
ID   Q4TTJ2_STRPU PRELIMINARY;      PRT;    39 AA.
AC   Q4TTJ2;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Extracellular protein (Fragment).
GN   Name=endo16;
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15937122;
RA   Balhoff J.P., Wray G.A.;
RT   "Evolutionary analysis of the well characterized endo16 promoter
RT   reveals substantial variation within functional sites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 102:8591-8596(2005).
DR   EMBL; DQ066806; AAY51735.1; -; Genomic_DNA.
DR   EMBL; DQ066809; AAY51738.1; -; Genomic_DNA.
DR   EMBL; DQ066811; AAY51740.1; -; Genomic_DNA.
DR   EMBL; DQ066812; AAY51741.1; -; Genomic_DNA.
DR   EMBL; DQ066813; AAY51742.1; -; Genomic_DNA.
DR   EMBL; DQ066816; AAY51745.1; -; Genomic_DNA.
DR   EMBL; DQ066801; AAY51730.1; -; Genomic_DNA.
FT   NON_TER      39     39
SQ   SEQUENCE   39 AA;  4351 MW;  2AA9F34D78BE902A CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 18
Q4TTJ4_STRPU
ID   Q4TTJ4_STRPU PRELIMINARY;      PRT;    39 AA.
AC   Q4TTJ4;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Extracellular protein (Fragment).
GN   Name=endo16;
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15937122;
RA   Balhoff J.P., Wray G.A.;
RT   "Evolutionary analysis of the well characterized endo16 promoter
RT   reveals substantial variation within functional sites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 102:8591-8596(2005).
DR   EMBL; DQ066804; AAY51733.1; -; Genomic_DNA.
FT   NON_TER      39     39
SQ   SEQUENCE   39 AA;  4337 MW;  7A9AF34D79FB9028 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 19
Q4TTJ8_STRPU
ID   Q4TTJ8_STRPU PRELIMINARY;      PRT;    39 AA.
AC   Q4TTJ8;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Extracellular protein (Fragment).
GN   Name=endo16;
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15937122;
```

```
RA   Balhoff J.P., Wray G.A.;
RT   "Evolutionary analysis of the well characterized endo16 promoter
RT   reveals substantial variation within functional sites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 102:8591-8596(2005).
DR   EMBL; DQ066800; AAY51729.1; -; Genomic_DNA.
DR   EMBL; DQ066802; AAY51731.1; -; Genomic_DNA.
DR   EMBL; DQ066803; AAY51732.1; -; Genomic_DNA.
DR   EMBL; DQ066810; AAY51739.1; -; Genomic_DNA.
DR   EMBL; DQ066818; AAY51747.1; -; Genomic_DNA.
DR   EMBL; DQ066799; AAY51728.1; -; Genomic_DNA.
FT   NON_TER      39     39
SQ   SEQUENCE   39 AA;  4365 MW;  2AA9F34D79FB902A CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 20
Q4TTK1_STRPU
ID   Q4TTK1_STRPU PRELIMINARY;      PRT;    39 AA.
AC   Q4TTK1;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Extracellular protein (Fragment).
GN   Name=endo16;
OS   Strongylocentrotus purpuratus (Purple sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7668;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15937122;
RA   Balhoff J.P., Wray G.A.;
RT   "Evolutionary analysis of the well characterized endo16 promoter
RT   reveals substantial variation within functional sites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 102:8591-8596(2005).
DR   EMBL; DQ066797; AAY51726.1; -; Genomic_DNA.
FT   NON_TER      39     39
SQ   SEQUENCE   39 AA;  4399 MW;  2AA9F34D78BE8B2A CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 21
Q66546_9GAMA
ID   Q66546_9GAMA PRELIMINARY;      PRT;    39 AA.
AC   Q66546;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   EBV B95-8 Cl(e) DNA with antigen coding ORF (Fragment).
OS   Human herpesvirus 4 (Epstein-Barr virus).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Lymphocryptovirus.
OX   NCBI_TaxID=10376;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   MEDLINE=88217505; PubMed=2835748;
RA   Walls D., Perricaudet M., Gannon F.;
RT   "The analysis of EBV proteins which are antigenic in vivo.";
RL   Nucleic Acids Res. 16:2859-2872(1988).
DR   EMBL; X07530; CAA30405.1; -; Genomic_DNA.
FT   NON_TER      39     39
SQ   SEQUENCE   39 AA;  4280 MW;  7937120F24AF5774 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 39;
  Best Local Similarity   100.0%;  Pred. No. 4.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         10 RRLN 13


RESULT 22
Q4TTG6_STRDR
ID   Q4TTG6_STRDR PRELIMINARY;      PRT;    40 AA.
AC   Q4TTG6;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Extracellular protein (Fragment).
GN   Name=endo16;
OS   Strongylocentrotus droebachiensis (Sea urchin).
OC   Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC   Echinoidea; Euechinoidea; Echinacea; Echinoida; Strongylocentrotidae;
OC   Strongylocentrotus.
OX   NCBI_TaxID=7671;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15937122;
RA   Balhoff J.P., Wray G.A.;
RT   "Evolutionary analysis of the well characterized endo16 promoter
RT   reveals substantial variation within functional sites.";
RL   Proc. Natl. Acad. Sci. U.S.A. 102:8591-8596(2005).
DR   EMBL; DQ066862; AAY52171.1; -; Genomic_DNA.
FT   NON_TER      40     40
SQ   SEQUENCE   40 AA;  4540 MW;  70855D302101C533 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 40;
  Best Local Similarity   100.0%;  Pred. No. 4.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 23
Q663L4_YERPS
ID   Q663L4_YERPS PRELIMINARY;      PRT;    40 AA.
AC   Q663L4;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=pYV0052;
OS   Yersinia pseudotuberculosis.
OG   Plasmid pYV.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Yersinia.
OX   NCBI_TaxID=633;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=IP32953 / Serotype I;
RX   PubMed=15358858; DOI=10.1073/pnas.0404012101;
RA   Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA   Regala W.M., Georgescu A.M., Vergez L.M., Land M.L., Motin V.L.,
RA   Brubaker R.R., Fowler J., Hinnebusch J., Marceau M., Medigue C.,
RA   Simonet M., Chenal-Francisque V., Souza B., Dacheux D., Elliott J.M.,
RA   Derbise A., Hauser L.J., Garcia E.;
RT   "Insights into the evolution of Yersinia pestis through whole-genome
RT   comparison with Yersinia pseudotuberculosis.";
```

```
RL   Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR   EMBL; BX936399; CAF25395.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein; Plasmid.
SQ   SEQUENCE   40 AA;  4820 MW;  550F7DF129673FB0 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 40;
  Best Local Similarity   100.0%;  Pred. No. 4.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         25 RRLN 28


RESULT 24
Q7RL48_PLAYO
ID   Q7RL48_PLAYO PRELIMINARY;      PRT;    43 AA.
AC   Q7RL48;
DT   01-MAR-2004 (TrEMBLrel. 26, Created)
DT   01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT   01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE   Hypothetical protein.
GN   Name=PY02698;
OS   Plasmodium yoelii yoelii.
OC   Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX   NCBI_TaxID=73239;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=17XNL;
RX   MEDLINE=22255706; PubMed=12368865; DOI=10.1038/nature01099;
RA   Carlton J.M., Angiuoli S.V., Suh B.B., Kooij T.W., Pertea M.,
RA   Silva J.C., Ermolaeva M.D., Allen J.E., Selengut J.D., Koo H.L.,
RA   Peterson J.D., Pop M., Kosack D.S., Shumway M.F., Bidwell S.L.,
RA   Shallom S.J., van Aken S.E., Riedmuller S.B., Feldblyum T.V.,
RA   Cho J.K., Quackenbush J., Sedegah M., Shoaibi A., Cummings L.M.,
RA   Florens L., Yates J.R. III, Raine J.D., Sinden R.E., Harris M.A.,
RA   Cunningham D.A., Preiser P.R., Bergman L.W., Vaidya A.B.,
RA   van Lin L.H., Janse C.J., Waters A.P., Smith H.O., White O.R.,
RA   Salzberg S.L., Venter J.C., Fraser C.M., Hoffman S.L., Gardner M.J.,
RA   Carucci D.J.;
RT   "Genome sequence and comparative analysis of the model rodent malaria
RT   parasite Plasmodium yoelii yoelii.";
RL   Nature 419:512-519(2002).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AABL01000743; EAA22179.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   43 AA;  5154 MW;  42B52DAD652E255B CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 43;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         15 RRLN 18


RESULT 25
Q8KEJ5_CHLTE
ID   Q8KEJ5_CHLTE PRELIMINARY;      PRT;    43 AA.
AC   Q8KEJ5;
DT   01-OCT-2002 (TrEMBLrel. 22, Created)
DT   01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=CT0693;
OS   Chlorobium tepidum.
OC   Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC   Chlorobaculum.
OX   NCBI_TaxID=1097;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=TLS / ATCC 49652 / DSM 12025;
RX   MEDLINE=22103685; PubMed=12093901; DOI=10.1073/pnas.132181499;
RA   Eisen J.A., Nelson K.E., Paulsen I.T., Heidelberg J.F., Wu M.,
RA   Dodson R.J., DeBoy R.T., Gwinn M.L., Nelson W.C., Haft D.H.,
RA   Hickey E.K., Peterson J.D., Durkin A.S., Kolonay J.F., Yang F.,
RA   Holt I.E., Umayam L.A., Mason T.M., Brenner M., Shea T.P.,
RA   Parksey D.S., Nierman W.C., Feldblyum T.V., Hansen C.L., Craven M.B.,
RA   Radune D., Vamathevan J.J., Khouri H.M., White O., Gruber T.M.,
RA   Ketchum K.A., Venter J.C., Tettelin H., Bryant D.A., Fraser C.M.;
RT   "The complete genome sequence of Chlorobium tepidum TLS, a
RT   photosynthetic, anaerobic, green-sulfur bacterium.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:9509-9514(2002).
DR   EMBL; AE006470; AAM71931.1; -; Genomic_DNA.
DR   TIGR; CT0693; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   43 AA;  5354 MW;  EF0A76D438C1C2B7 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 43;
  Best Local Similarity   100.0%;  Pred. No. 5.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          8 RRLN 11


RESULT 26
Q5C5H2_SCHJA
ID   Q5C5H2_SCHJA PRELIMINARY;      PRT;    44 AA.
AC   Q5C5H2;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein.
OS   Schistosoma japonicum (Blood fluke).
OC   Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida;
OC   Schistosomatoidea; Schistosomatidae; Schistosoma.
OX   NCBI_TaxID=6182;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Han Z.;
RL   Submitted (MAR-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AY809213; AAX25102.1; -; mRNA.
DR   InterPro; IPR008919; Retrov_capsid_N.
KW   Hypothetical protein.
SQ   SEQUENCE   44 AA;  5441 MW;  A96DEF5A34DAF9F1 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         34 RRLN 37


RESULT 27
Q72AJ8_DESVH
ID   Q72AJ8_DESVH PRELIMINARY;      PRT;    44 AA.
AC   Q72AJ8;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=DVU1994;
OS   Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS   8303).
OC   Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC   Desulfovibrionaceae; Desulfovibrio.
OX   NCBI_TaxID=882;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
```

```
RX   PubMed=15077118; DOI=10.1038/nbt959;
RA   Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA   Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA   Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA   Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA   Peterson J.D., Davidsen T.M., Zafar N., Zhou L., Radune D.,
RA   Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA   Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT   "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT   Desulfovibrio vulgaris Hildenborough.";
RL   Nat. Biotechnol. 22:554-559(2004).
DR   EMBL; AE017314; AAS96470.1; -; Genomic_DNA.
DR   TIGR; DVU1994; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   44 AA;  5059 MW;  EF626D0A31D46E3C CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 44;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         33 RRLN 36


RESULT 28
Q99IR3_9GEMI
ID   Q99IR3_9GEMI PRELIMINARY;      PRT;    47 AA.
AC   Q99IR3;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Coat protein (Fragment).
OS   Tomato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28350;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Samretwanich K., Chiemsombat P., Kittipakorn K., Ikegami M.;
RT   "Yellow leaf disease of muskmelon from Thailand caused by Tomato leaf
RT   curl virus.";
RL   Plant Dis. 84:707-707(2000).
DR   EMBL; AB020976; BAB33262.1; -; Genomic_DNA.
KW   Capsid protein.
FT   NON_TER      47     47
SQ   SEQUENCE   47 AA;  5281 MW;  60766BA44199E0C8 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 47;
  Best Local Similarity   100.0%;  Pred. No. 5.8e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         19 RRLN 22


RESULT 29
Q99IR5_9GEMI
ID   Q99IR5_9GEMI PRELIMINARY;      PRT;    47 AA.
AC   Q99IR5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   Coat protein (Fragment).
OS   Tomato leaf curl virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=28350;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Samretwanich K., Chiemsombat P., Kittipakorn K., Ikegami M.;
RT   "Tomato leaf curl geminivirus assiciated with cucumber yellow disease
RT   in Thailand.";
RL   J. Phytopathol. 148:615-617(2000).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RA   Samretwanich K., Chiemsombat P., Kittipakorn K., Ikegami M.;
RT   "Yellow disease of cantaloupe and wax gourd from Thailand caused by
RT   Tomato leaf curl virus.";
RL   Plant Dis. 84:200-200(2000).
RN   [3]
RP   NUCLEOTIDE SEQUENCE.
RA   Samretwanich K., Chiemsombat P., Kittipakorn K., Ikegami M.;
RT   "Tomato leaf curl geminivirus associated with cantaloupe yellow leaf
RT   disease in Thailand.";
RL   World J. Microbiol. Biotechnol. 16:401-403(2000).
DR   EMBL; AB017342; BAB33260.1; -; Genomic_DNA.
KW   Capsid protein.
FT   NON_TER      47     47
SQ   SEQUENCE   47 AA;  5281 MW;  60766BA44199E0C8 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 47;
  Best Local Similarity   100.0%;  Pred. No. 5.8e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         19 RRLN 22


RESULT 30
Q99IR1_9GEMI
ID   Q99IR1_9GEMI PRELIMINARY;      PRT;    48 AA.
AC   Q99IR1;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   Coat protein (Fragment).
OS   Soybean crinkle leaf virus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=85753;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Samretwanich K., Kittipakorn K., Chiemsombat P., Ikegami M.;
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB020977; BAB33264.1; -; Genomic_DNA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000650; Gem_coat_AR1.
DR   ProDom; PD000901; Gem_coat_AR1; 1.
KW   Capsid protein.
FT   NON_TER      48     48
SQ   SEQUENCE   48 AA;  5509 MW;  200765C58F2FDBA8 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 48;
  Best Local Similarity   100.0%;  Pred. No. 5.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         19 RRLN 22


RESULT 31
PRTZ1_SCYCA
ID   PRTZ1_SCYCA    STANDARD;      PRT;    50 AA.
AC   P08433;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-NOV-1988 (Rel. 09, Last sequence update)
DT   10-MAY-2005 (Rel. 47, Last annotation update)
DE   Protamine Z1 (Scylliorhinine Z1).
OS   Scyliorhinus canicula (Spotted dogfish) (Spotted catshark).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Chondrichthyes;
OC   Elasmobranchii; Galeomorphii; Galeoidea; Carcharhiniformes;
OC   Scyliorhinidae; Scyliorhinus.
OX   NCBI_TaxID=7830;
RN   [1]
```

```
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Testis;
RX   MEDLINE=87246639; PubMed=3595601;
RA   Berlot-Picard F., Vodjdani G., Doly J.;
RT   "Isolation and characterization of a cDNA clone encoding testis
RT   protamine Z1 from the dog-fish Scylliorhinus caniculus.";
RL   Eur. J. Biochem. 165:553-557(1987).
RN   [2]
RP   PROTEIN SEQUENCE.
RX   MEDLINE=84000513; PubMed=6615852; DOI=10.1016/0167-4838(83)90031-6;
RA   Gusse M., Sautiere P., Chauviere M., Chevaillier P.;
RT   "Extraction, purification and characterization of the sperm protamines
RT   of the dog-fish Scylliorhinus caniculus.";
RL   Biochim. Biophys. Acta 748:93-98(1983).
CC   -!- FUNCTION: Protamines substitute for histones in the chromatin of
CC       sperm during the haploid phase of spermatogenesis. They compact
CC       sperm DNA into a highly condensed, stable and inactive complex.
CC   -!- SUBCELLULAR LOCATION: Nuclear.
CC   -!- TISSUE SPECIFICITY: Testis.
CC   ---------------------------------------------------------------------
CC   This Swiss-Prot entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use as long as its content is in no way modified and this statement is not
CC   removed.
CC   ---------------------------------------------------------------------
DR   EMBL; X05611; CAA29099.1; -; mRNA.
DR   PIR; S00016; S00016.
KW   Chromosomal protein; Developmental protein; Differentiation;
KW   Direct protein sequencing; DNA condensation; DNA-binding;
KW   Nuclear protein; Nucleosome core; Spermatogenesis; Testis.
FT   INIT_MET      0      0
SQ   SEQUENCE   50 AA;  6344 MW;  0C6BDEA9D128BDB8 CRC64;

  Query Match             95.2%;  Score 20;  DB 1;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         20 RRLN 23


RESULT 32
Q6DTN0_CANGA
ID   Q6DTN0_CANGA PRELIMINARY;      PRT;    50 AA.
AC   Q6DTN0;
DT   25-OCT-2004 (TrEMBLrel. 28, Created)
DT   25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT   25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE   Pps1p (Fragment).
GN   Name=PPS1;
OS   Candida glabrata (Yeast) (Torulopsis glabrata).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX   NCBI_TaxID=5478;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RX   PubMed=15686568; DOI=10.1111/j.1365-2958.2004.04465.x;
RA   Castano I., Pan S.J., Zupancic M., Hennequin C., Dujon B.,
RA   Cormack B.P.;
RT   "Telomere length control and transcriptional regulation of
RT   subtelomeric adhesins in Candida glabrata.";
RL   Mol. Microbiol. 55:1246-1258(2005).
DR   EMBL; AY646924; AAT67385.1; -; Genomic_DNA.
FT   NON_TER       1      1
SQ   SEQUENCE   50 AA;  6269 MW;  FF14B5529209E623 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 33
Q831H1_ENTFA
ID   Q831H1_ENTFA PRELIMINARY;      PRT;    50 AA.
AC   Q831H1;
DT   01-JUN-2003 (TrEMBLrel. 24, Created)
DT   01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=EF2538;
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=V583 / ATCC 700802;
RX   MEDLINE=22550857; PubMed=12663927; DOI=10.1126/science.1080613;
RA   Paulsen I.T., Banerjei L., Myers G.S.A., Nelson K.E., Seshadri R.,
RA   Read T.D., Fouts D.E., Eisen J.A., Gill S.R., Heidelberg J.F.,
RA   Tettelin H., Dodson R.J., Umayam L.A., Brinkac L.M., Beanan M.J.,
RA   Daugherty S.C., DeBoy R.T., Durkin S.A., Kolonay J.F., Madupu R.,
RA   Nelson W.C., Vamathevan J.J., Tran B., Upton J., Hansen T., Shetty J.,
RA   Khouri H.M., Utterback T.R., Radune D., Ketchum K.A., Dougherty B.A.,
RA   Fraser C.M.;
RT   "Role of mobile DNA in the evolution of vancomycin-resistant
RT   Enterococcus faecalis.";
RL   Science 299:2071-2074(2003).
DR   EMBL; AE016955; AAO82251.1; -; Genomic_DNA.
DR   TIGR; EF2538; -.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   50 AA;  6083 MW;  9DB8BE43B9A11352 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         18 RRLN 21


RESULT 34
Q4T9A7_TETNG
ID   Q4T9A7_TETNG PRELIMINARY;      PRT;    50 AA.
AC   Q4T9A7;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Chromosome 3 SCAF7608, whole genome shotgun sequence. (Fragment).
GN   ORFNames=GSTENG00004841001;
OS   Tetraodon nigroviridis (Green puffer).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
OC   Tetradontoidea; Tetraodontidae; Tetraodon.
OX   NCBI_TaxID=99883;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Jaillon O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA   Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA   Nicaud S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA   Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA   Anthouard V., Jubin C., Castelli V., Katinka M., Vacherie B.,
RA   Biemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,
RA   Cruaud C., Duprat S., Brottier P., Coutanceau J.P., Gouzy J.,
RA   Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA   Kellis M., Volff JN., Guigo R., Zody M.C., Mesirov J.,
RA   Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA   Laudet V., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA   Wincker P., Lander E.S., Weissenbach J., Roest Crollius H.;
```

```
RT   "Genome duplication in the teleost fish Tetraodon nigroviridis reveals
RT   the early vertebrate proto-karyotype.";
RL   Nature 431:946-957(2004).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RG   Genoscope; Whitehead Institute Centre for Genome Research;
RL   Submitted (FEB-2004) to the EMBL/GenBank/DDBJ databases.
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; CAAE01007608; CAF90525.1; -; Genomic_DNA.
FT   NON_TER       1      1
FT   NON_TER      50     50
SQ   SEQUENCE   50 AA;  6133 MW;  F6C79EC009632BDC CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 50;
  Best Local Similarity   100.0%;  Pred. No. 6.2e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         42 RRLN 45


RESULT 35
Q9ZSF6_9SOLA
ID   Q9ZSF6_9SOLA PRELIMINARY;      PRT;    51 AA.
AC   Q9ZSF6;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Self-incompatibility ribonuclease S15 (Fragment).
OS   Lycium andersonii.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   asterids; lamiids; Solanales; Solanaceae; Lycium.
OX   NCBI_TaxID=87540;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Richman A.D., Kohn J.R.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF105357; AAD13085.1; -; mRNA.
DR   GO; GO:0004521; F:endoribonuclease activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
DR   InterPro; IPR001568; RNase_T2.
DR   Pfam; PF00445; Ribonuclease_T2; 1.
FT   NON_TER       1      1
FT   NON_TER      51     51
SQ   SEQUENCE   51 AA;  6289 MW;  C6A4248AC15ABA9E CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          4 RRLN 7


RESULT 36
Q8RTK9_LEUME
ID   Q8RTK9_LEUME PRELIMINARY;      PRT;    51 AA.
AC   Q8RTK9;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Hypothetical protein.
OS   Leuconostoc mesenteroides.
OC   Bacteria; Firmicutes; Lactobacillales; Leuconostoc.
OX   NCBI_TaxID=1245;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Lee K.-H., Lee H.-J., Chang H.-C., Chung D.-K., Lee J.-H., Kim J.-H.;
RL   Submitted (SEP-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF420260; AAL77873.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   51 AA;  6177 MW;  A929B57F6D8907D9 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         14 RRLN 17


RESULT 37
Q8U4Z9_AGRT5
ID   Q8U4Z9_AGRT5 PRELIMINARY;      PRT;    51 AA.
AC   Q8U4Z9;
DT   01-JUN-2002 (TrEMBLrel. 21, Created)
DT   01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   AGR_L_771p.
GN   OrderedLocusNames=AGR_L_771;
OS   Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC   Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC   Rhizobiaceae; Rhizobium/Agrobacterium group; Agrobacterium.
OX   NCBI_TaxID=176299;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Cereon;
RX   MEDLINE=21608551; PubMed=11743194; DOI=10.1126/science.1066803;
RA   Goodner B., Hinkle G., Gattung S., Miller N., Blanchard M.,
RA   Qurollo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA   Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA   Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markelz B.,
RA   Flanagan C., Crowell C., Gurson J., Lomo C., Sear C., Strub G.,
RA   Cielo C., Slater S.;
RT   "Genome sequence of the plant pathogen and biotechnology agent
RT   Agrobacterium tumefaciens C58.";
RL   Science 294:2323-2328(2001).
DR   EMBL; AE008238; AAK88958.1; -; Genomic_DNA.
DR   PIR; D98179; D98179.
SQ   SEQUENCE   51 AA;  5890 MW;  918E0B5D948C97DF CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         34 RRLN 37


RESULT 38
Q57N53_SALCH
ID   Q57N53_SALCH PRELIMINARY;      PRT;    51 AA.
AC   Q57N53;
DT   10-MAY-2005 (TrEMBLrel. 30, Created)
DT   10-MAY-2005 (TrEMBLrel. 30, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Putative inner membrane protein.
GN   OrderedLocusNames=SC1952;
OS   Salmonella cholerae-suis (Salmonella enterica).
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=591;
RN   [1]
RP   NUCLEOTIDE SEQUENCE [LARGE SCALE GENOMIC DNA].
RC   STRAIN=SC-B67;
RX   PubMed=15781495;
RA   Chiu C.-H., Tang P., Chu C., Hu S., Bao Q., Yu J., Chou Y.-Y.,
RA   Wang H.-S., Lee Y.-S.;
```

```
RT   "The genome sequence of Salmonella enterica serovar Choleraesuis, a
RT   highly invasive and resistant zoonotic pathogen.";
RL   Nucleic Acids Res. 33:1690-1698(2005).
DR   EMBL; AE017220; AAX65858.1; -; Genomic_DNA.
KW   Complete proteome.
SQ   SEQUENCE   51 AA;  6024 MW;  61270A2A7D28DD76 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         33 RRLN 36


RESULT 39
Q8VKJ1_MYCTU
ID   Q8VKJ1_MYCTU PRELIMINARY;      PRT;    51 AA.
AC   Q8VKJ1;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=MT0576;
OS   Mycobacterium tuberculosis.
OC   Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC   Corynebacterineae; Mycobacteriaceae; Mycobacterium;
OC   Mycobacterium tuberculosis complex.
OX   NCBI_TaxID=1773;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CDC 1551 / Oshkosh;
RX   MEDLINE=22206494; PubMed=12218036;
RX   DOI=10.1128/JB.184.19.5479-5490.2002;
RA   Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA   Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA   Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA   Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA   Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA   Fraser C.M.;
RT   "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT   laboratory strains.";
RL   J. Bacteriol. 184:5479-5490(2002).
DR   EMBL; AE000516; AAK44799.1; -; Genomic_DNA.
DR   TIGR; MT0576; -.
KW   Hypothetical protein.
SQ   SEQUENCE   51 AA;  5489 MW;  B0803ECA77E3640B CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          2 RRLN 5


RESULT 40
Q5PI12_SALPA
ID   Q5PI12_SALPA PRELIMINARY;      PRT;    51 AA.
AC   Q5PI12;
DT   01-FEB-2005 (TrEMBLrel. 29, Created)
DT   01-FEB-2005 (TrEMBLrel. 29, Last sequence update)
DT   01-FEB-2005 (TrEMBLrel. 29, Last annotation update)
DE   Hypothetical protein.
GN   OrderedLocusNames=SPA0921;
OS   Salmonella paratyphi-a.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=54388;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=ATCC 9150;
RX   PubMed=15531882; DOI=10.1038/ng1470;
RA   McClelland M., Sanderson K.E., Clifton S.W., Latreille P.,
RA   Porwollik S., Sabo A., Meyer R., Bieri T., Ozersky P., McLellan M.,
RA   Harkins C.R., Wang C., Nguyen C., Berghoff A., Elliott G.,
RA   Kohlberg S., Strong C., Du F., Carter J., Kremizki C., Layman D.,
RA   Leonard S., Sun H., Fulton L., Nash W., Miner T., Minx P.,
RA   Delehaunty K., Fronick C., Magrini V., Nhan M., Warren W., Florea L.,
RA   Spieth J., Wilson R.K.;
RT   "Comparison of genome degradation in Paratyphi A and Typhi, human-
RT   restricted serovars of Salmonella enterica that cause typhoid.";
RL   Nat. Genet. 36:1268-1274(2004).
DR   EMBL; CP000026; AAV76900.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   51 AA;  6011 MW;  EC78D82A7D398933 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         33 RRLN 36


RESULT 41
Q8Z5T2_SALTI
ID   Q8Z5T2_SALTI PRELIMINARY;      PRT;    51 AA.
AC   Q8Z5T2; Q7CAN6;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   10-MAY-2005 (TrEMBLrel. 30, Last annotation update)
DE   Hypothetical protein STY2156.
GN   OrderedLocusNames=STY2156, t0928;
OS   Salmonella typhi.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=601;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=CT18;
RX   MEDLINE=21534947; PubMed=11677608; DOI=10.1038/35101607;
RA   Parkhill J., Dougan G., James K.D., Thomson N.R., Pickard D., Wain J.,
RA   Churcher C.M., Mungall K.L., Bentley S.D., Holden M.T.G., Sebaihia M.,
RA   Baker S., Basham D., Brooks K., Chillingworth T., Connerton P.,
RA   Cronin A., Davis P., Davies R.M., Dowd L., White N., Farrar J.,
RA   Feltwell T., Hamlin N., Haque A., Hien T.T., Holroyd S., Jagels K.,
RA   Krogh A., Larsen T.S., Leather S., Moule S., O'Gaora P., Parry C.,
RA   Quail M.A., Rutherford K.M., Simmonds M., Skelton J., Stevens K.,
RA   Whitehead S., Barrell B.G.;
RT   "Complete genome sequence of a multiple drug resistant Salmonella
RT   enterica serovar Typhi CT18.";
RL   Nature 413:848-852(2001).
RN   [2]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=Ty2 / ATCC 700931;
RX   MEDLINE=22531367; PubMed=12644504;
RX   DOI=10.1128/JB.185.7.2330-2337.2003;
RA   Deng W., Liou S.-R., Plunkett G. III, Mayhew G.F., Rose D.J.,
RA   Burland V., Kodoyianni V., Schwartz D.C., Blattner F.R.;
RT   "Comparative genomics of Salmonella enterica serovar Typhi strains Ty2
RT   and CT18.";
RL   J. Bacteriol. 185:2330-2337(2003).
DR   EMBL; AL627272; CAD05697.1; -; Genomic_DNA.
DR   EMBL; AE016837; AAO68605.1; -; Genomic_DNA.
KW   Complete proteome; Hypothetical protein.
SQ   SEQUENCE   51 AA;  6011 MW;  EC78D82A7D398933 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
```

```
              ||||
Db         33 RRLN 36


RESULT 42
Q8ZNT9_SALTY
ID   Q8ZNT9_SALTY PRELIMINARY;      PRT;    51 AA.
AC   Q8ZNT9;
DT   01-MAR-2002 (TrEMBLrel. 20, Created)
DT   01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Putative inner membrane protein.
GN   OrderedLocusNames=STM1948;
OS   Salmonella typhimurium.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC   Enterobacteriaceae; Salmonella.
OX   NCBI_TaxID=602;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=LT2;
RX   MEDLINE=21534948; PubMed=11677609; DOI=10.1038/35101614;
RA   McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA   Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA   Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA   Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA   Waterston R., Wilson R.K.;
RT   "Complete genome sequence of Salmonella enterica serovar Typhimurium
RT   LT2.";
RL   Nature 413:852-856(2001).
DR   EMBL; AE008786; AAL20860.1; -; Genomic_DNA.
KW   Complete proteome.
SQ   SEQUENCE   51 AA;  5995 MW;  8078D82A7D28CC27 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 51;
  Best Local Similarity   100.0%;  Pred. No. 6.3e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         33 RRLN 36


RESULT 43
Q96I42_HUMAN
ID   Q96I42_HUMAN PRELIMINARY;      PRT;    52 AA.
AC   Q96I42;
DT   01-DEC-2001 (TrEMBLrel. 19, Created)
DT   01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT   01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE   Hypothetical protein (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC   Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (MAY-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC007841; AAH07841.1; -; mRNA.
KW   Hypothetical protein.
FT   NON_TER       1      1
SQ   SEQUENCE   52 AA;  5746 MW;  13C8F9B7A0532365 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         38 RRLN 41


RESULT 44
Q54BT9_DICDI
ID   Q54BT9_DICDI PRELIMINARY;      PRT;    52 AA.
AC   Q54BT9;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Hypothetical protein.
GN   ORFNames=DDB0191934;
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=AX4;
RA   Eichinger L., Pachebat J.A., Gloeckner G., Rajandream M.-A.,
RA   Sucgang R., Berriman M., Song J., Olsen R., Szafranski K., Xu Q.,
RA   Tunggal B., Kummerfeld S., Madera M., Konfortov B.A., Rivero F.,
RA   Bankier A.T., Lehmann R., Hamlin N., Davies R., Gaudet P., Fey P.,
RA   Pilcher K., Chen G., Saunders D., Sodergren E., Davis P.,
RA   Kerhornou A., Nie X., Hall N., Anjard C., Hemphill L., Bason N.,
RA   Farbrother P., Desany B., Just E., Morio T., Rost R., Churcher C.,
RA   Cooper J., Haydock S., van Driessche N., Cronin A., Goodhead I.,
RA   Muzny D., Mourier T., Pain A., Lu M., Harper D., Lindsay R.,
RA   Hauser H., James K., Quiles M., Mohan M.B., Saito T., Buchrieser C.,
RA   Wardroper A., Felder M., Thangavelu M., Johnson D., Knights A.,
RA   Loulseged H., Mungall K., Oliver K., Price C., Quail M.A.,
RA   Urushihara H., Hernandez J., Rabbinowitsch E., Steffen D., Sanders M.,
RA   Ma J., Kohara Y., Sharp S., Simmonds M., Spiegler S., Tivey A.,
RA   Sugano S., White B., Walker D., Woodward J., Winckler T., Tanaka Y.,
RA   Shaulsky G., Schleicher M., Weinstock G., Rosenthal A., Cox E.C.,
RA   Chisholm R.L., Gibbs R., Loomis W.F., Platzer M., Kay R.R.,
RA   Williams J., Dear P.H., Noegel A.A., Barrell B., Kuspa A.;
RT   "The genome of the social amoeba Dictyostelium discoideum.";
RL   Nature 0:0-0(2005).
CC   -!- CAUTION: The sequence shown here is derived from an
CC       EMBL/GenBank/DDBJ whole genome shotgun (WGS) entry which is
CC       preliminary data.
DR   EMBL; AAFI01000284; EAL60727.1; -; Genomic_DNA.
KW   Hypothetical protein.
SQ   SEQUENCE   52 AA;  6226 MW;  1C226192CE0CE606 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 52;
  Best Local Similarity   100.0%;  Pred. No. 6.5e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         19 RRLN 22


RESULT 45
Q9ZSG7_9SOLA
ID   Q9ZSG7_9SOLA PRELIMINARY;      PRT;    53 AA.
AC   Q9ZSG7;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE   Self-incompatibility ribonuclease S4 (Fragment).
OS   Lycium andersonii.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
OC   asterids; lamiids; Solanales; Solanaceae; Lycium.
OX   NCBI_TaxID=87540;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Richman A.D., Kohn J.R.;
RL   Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF105346; AAD13074.1; -; mRNA.
DR   GO; GO:0004521; F:endoribonuclease activity; IEA.
DR   GO; GO:0016787; F:hydrolase activity; IEA.
DR   GO; GO:0003723; F:RNA binding; IEA.
```

```
DR   InterPro; IPR001568; RNase_T2.
DR   Pfam; PF00445; Ribonuclease_T2; 1.
FT   NON_TER       1      1
FT   NON_TER      53     53
SQ   SEQUENCE   53 AA;  6470 MW;  9ACDE8FE804A414E CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 53;
  Best Local Similarity   100.0%;  Pred. No. 6.6e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db          4 RRLN 7


RESULT 46
Q8EDQ1_SHEON
ID   Q8EDQ1_SHEON PRELIMINARY;      PRT;    54 AA.
AC   Q8EDQ1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein SO2692.
GN   OrderedLocusNames=SO2692;
OS   Shewanella oneidensis.
OC   Bacteria; Proteobacteria; Gammaproteobacteria; Alteromonadales;
OC   Shewanellaceae; Shewanella.
OX   NCBI_TaxID=70863;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=MR-1;
RX   MEDLINE=22297686; PubMed=12368813; DOI=10.1038/nbt749;
RA   Heidelberg J.F., Paulsen I.T., Nelson K.E., Gaidos E.J., Nelson W.C.,
RA   Read T.D., Eisen J.A., Seshadri R., Ward N.L., Methe B.A.,
RA   Clayton R.A., Meyer T., Tsapin A., Scott J., Beanan M.J.,
RA   Brinkac L.M., Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S.,
RA   Haft D.H., Kolonay J.F., Madupu R., Peterson J.D., Umayam L.A.,
RA   White O., Wolf A.M., Vamathevan J.J., Weidman J.F., Impraim M.,
RA   Lee K., Berry K.J., Lee C., Mueller J., Khouri H.M., Gill J.,
RA   Utterback T.R., McDonald L.A., Feldblyum T.V., Smith H.O.,
RA   Venter J.C., Nealson K.H., Fraser C.M.;
RT   "Genome sequence of the dissimilatory metal ion-reducing bacterium
RT   Shewanella oneidensis.";
RL   Nat. Biotechnol. 20:1118-1123(2002).
DR   EMBL; AE015709; AAN55720.1; -; Genomic_DNA.
DR   TIGR; SO2692; -.
KW   Complete proteome.
SQ   SEQUENCE   54 AA;  5964 MW;  8AFC3DAB6DBA2D37 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 54;
  Best Local Similarity   100.0%;  Pred. No. 6.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         37 RRLN 40


RESULT 47
Q8DI57_SYNEL
ID   Q8DI57_SYNEL PRELIMINARY;      PRT;    54 AA.
AC   Q8DI57;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Tsr1734 protein.
GN   OrderedLocusNames=tsr1734;
OS   Synechococcus elongatus (Thermosynechococcus elongatus).
OC   Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX   NCBI_TaxID=32046;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RC   STRAIN=BP-1;
RX   MEDLINE=22225144; PubMed=12240834;
RA   Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA   Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA   Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA   Shimpo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT   "Complete genome structure of the thermophilic cyanobacterium
RT   Thermosynechococcus elongatus BP-1.";
RL   DNA Res. 9:123-130(2002).
DR   EMBL; BA000039; BAC09286.1; -; Genomic_DNA.
KW   Complete proteome.
SQ   SEQUENCE   54 AA;  6487 MW;  2F9494243FB80E67 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 54;
  Best Local Similarity   100.0%;  Pred. No. 6.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         32 RRLN 35


RESULT 48
Q9E3W1_TYLCV
ID   Q9E3W1_TYLCV PRELIMINARY;      PRT;    54 AA.
AC   Q9E3W1;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE   Coat protein (Fragment).
OS   Tomato yellow leaf curl virus (TYLCV).
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus.
OX   NCBI_TaxID=10832;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Ying Z., Davis M.J.;
RT   "Partial characterization and host range of tomato yellow leaf curl
RT   virus in south Florida.";
RL   Annu. Meet. Fla. State Hort. Soc. 0:0-0(2000).
DR   EMBL; AF260331; AAG23626.1; -; Genomic_DNA.
DR   GO; GO:0019028; C:viral capsid; IEA.
DR   GO; GO:0005198; F:structural molecule activity; IEA.
DR   InterPro; IPR000650; Gem_coat_AR1.
DR   ProDom; PD000901; Gem_coat_AR1; 1.
KW   Capsid protein.
FT   NON_TER      54     54
SQ   SEQUENCE   54 AA;  6349 MW;  A626B63B92E07376 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 54;
  Best Local Similarity   100.0%;  Pred. No. 6.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         19 RRLN 22


RESULT 49
Q4GXV3_9GEMI
ID   Q4GXV3_9GEMI PRELIMINARY;      PRT;    54 AA.
AC   Q4GXV3;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Coat protein (Fragment).
GN   Name=AV1;
OS   Tomato geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus;
OC   unclassified Begomovirus.
OX   NCBI_TaxID=32610;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Li Z., Zhou X.;
RT   "Molecular characterization of tomato-infecting gomoviruses in Yunnan,
```

```
RT   China.";
RL   Submitted (JUL-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AM048873; CAJ15534.1; -; Genomic_DNA.
KW   Capsid protein.
FT   NON_TER      54     54
SQ   SEQUENCE   54 AA;  6249 MW;  23B2A35D6D8759F0 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 54;
  Best Local Similarity   100.0%;  Pred. No. 6.7e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         19 RRLN 22


RESULT 50
Q4GXV7_9GEMI
ID   Q4GXV7_9GEMI PRELIMINARY;      PRT;    55 AA.
AC   Q4GXV7;
DT   13-SEP-2005 (TrEMBLrel. 31, Created)
DT   13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT   13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE   Coat protein (Fragment).
GN   Name=AV1;
OS   Tomato geminivirus.
OC   Viruses; ssDNA viruses; Geminiviridae; Begomovirus;
OC   unclassified Begomovirus.
OX   NCBI_TaxID=32610;
RN   [1]
RP   NUCLEOTIDE SEQUENCE.
RA   Li Z., Zhou X.;
RT   "Molecular characterization of tomato-infecting gomoviruses in Yunnan,
RT   China.";
RL   Submitted (JUL-2005) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AM048871; CAJ15530.1; -; Genomic_DNA.
KW   Capsid protein.
FT   NON_TER      55     55
SQ   SEQUENCE   55 AA;  6324 MW;  70A2D5A34EF03191 CRC64;

  Query Match             95.2%;  Score 20;  DB 2;  Length 55;
  Best Local Similarity   100.0%;  Pred. No. 6.9e+02;
  Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 RRLN 4
              ||||
Db         19 RRLN 22


Search completed: January 25, 2006, 18:41:54
Job time : 103 secs
```